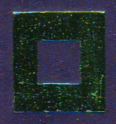
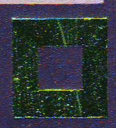
ВЫСШЕЕ ОБРАЗОВАНИЕ

СОВРЕМЕННЫЙ УЧЕБНИК



Л.Д. Кудрявцев



КУРС МАТЕМАТИЧЕСКОГО АНАЛИЗА

1

дрофа

Предисловие

В предлагаемом курсе математического анализа излагаются как традиционные классические методы, так и современные, которые возникли в последние десятилетия. Действительные числа вводятся аксиоматически. Этот путь дает возможность наиболее компактно и полно изложить необходимые для анализа сведения о числах. Вместе с тем он и логически наиболее совершенен, поскольку при других, так называемых «конструктивных», методах построения теории действительных чисел (когда за основу берутся бесконечные десятичные дроби, или сечения в области рациональных чисел, или классы эквивалентных фундаментальных последовательностей рациональных чисел) все равно необходимо вводить аксиому существования (непротиворечивости) множества действительных чисел, без которых проводимые построения не имеют логически завершенного характера. Поэтому проще всего сразу, исходя из аксиоматического задания действительных чисел, перейти к изучению математического анализа в собственном смысле слова.

В основном изложение материала в курсе ведется индуктивным методом: по возможности все вводимые понятия изучаются сначала в простейших ситуациях и лишь после обстоятельного их рассмотрения и накопления достаточного числа конкретных примеров производятся дальнейшие обобщения. Так, например, понятие предела сначала изучается для числовых последовательностей, затем для функций одного действительного переменного, далее вводится понятие предела отображения по множеству в евклидовом

пространстве, предела интегральных сумм и, наконец, все завершается рассмотрением общего понятия предела по фильтру в топологическом пространстве. Формула Тейлора рассматривается сначала для действительнозначной функции на отрезке, а в конце курса — для отображений линейных нормированных пространств: рассмотрение многочисленных критериев Коши существования тех или иных пределов завершается критерием Коши существования предела по фильтру отображения в полное метрическое пространство; изучение рядов Фурье начинается с рассмотрения классических тригонометрических рядов и заканчивается изучением рядов Фурье в гильбертовых пространствах по ортогональным системам; свойства непрерывных функций на отрезках обобщаются на отображения метрических компактов и континуумов и т. д.

Доказываемые теоремы не всегда формулируются с наибольшей общностью; иногда для лучшего выявления сущности изучаемого вопроса и идеи излагаемого доказательства рассмотрение проводится лишь для достаточно гладких функций. Такая точка зрения оправдана также тем, что, в силу плотности гладких функций в соответствующих функциональных пространствах, многие теоремы, доказанные для гладких функций, могут быть единым методом с помощью предельного перехода перенесены на более широкие классы функций. К сожалению, эту идею невозможно довести до конца без существенного увеличения объема книги. Поэтому вопрос о плотности «хороших» функций в различных функциональных пространствах рассмотрен в курсе лишь в простейших случаях.

Из существенных методических новшеств, которые автор считает целесообразными, следует отметить, что при определении предела функции по множеству при $x \to x_0$ не требуется выполнения условия $x \neq x_0$, так как это позволяет излагать вопросы, связанные с теорией пределов, проще и короче: например, само определение предела делается короче (на одно условие меньше), а тем самым упрощаются и доказательства, в которых участвуют пределы функций; не нужно рассматривать отдельно от теории пределов функций

непрерывные функции; делается наглядным и убедительным утверждение, что в математике дискретность является частным случаем непрерывности; упрощаются формулировка и доказательство важной теоремы о пределе композиции функций и т. д.

Основная часть учебника посвящена главным понятиям математического анализа и их свойствам. Овладев этим материалом, можно без особого труда самостоятельно справиться с изучением специальных вопросов математики, которые потребуются в дальнейшем по роду профессиональной деятельности.

Большое внимание в учебнике уделяется решению задач методами, основанными на излагаемой теории; кроме того, для самостоятельной работы рекомендуются упражнения и задачи. Решение упражнений весьма полезно для активного усвоения математического анализа. Вместе с тем их набор по своей полноте ни в коей мере не заменяет задачника. Отдельные же из предлагаемых задач довольно трудны. Их решение не является необходимым для овладения материалом и может потребовать довольно длительного времени. Как правило, они связаны с интересными и достаточно глубокими математическими фактами, подробное изложение которых ограничено объемом книги. Упражнения нумеруются отдельно в каждом параграфе, нумерация задач и рисунков — сквозная.

Изложение математического анализа ведется на уровне, доступном широкому кругу студентов. Вопросы, выходящие за рамки программ по высшей математике для инженерных и экономических специальностей и посвященные более глубокому изучению разделов анализа, необходимых студентам физико-математических специальностей, отмечены звездочкой. В силу этого учебник можно использовать в высших учебных заведениях с разным уровнем математического образования. Значительная часть материала, вошедшего в книгу, в течение многих лет читается автором в Московском физико-техническом институте в лекционном курсе математического анализа.

Настоящий учебник представляет собой переработанный трехтомный учебник автора «Курс математического анализа», М.: Высшая школа, 1988—1989.

Автор выражает глубокую благодарность своим коллегам по кафедре высшей математики С. М. Никольскому, В. С. Владимирову, О. В. Бесову, С. А. Теляковскому и Г. Н. Яковлеву, результаты многолетних обсуждений с которыми различных аспектов курса нашли свое непосредственное отражение в содержании книги.

Особенно автор признателен рецензентам книги — Н. В. Ефимову и В. А. Ильину, подробные и обстоятельные рецензии которых позволили во многом улучшить изложение материала.

Автор считает своим приятным долгом выразить благодарность преподавателям кафедры высшей математики Московского физико-технического института К. А. Бежанову, И. А. Борачинскому, Ф. Г. Булаевской, А. Д. Кутасову, В. А. Ходакову, В. И. Чехлову, сделавшим много полезных замечаний, которые были учтены при окончательном редактировании текста.

Введение

> Ни одно человеческое исследование не может называться истинной наукой, если оно не прошло через математические доказательства. Никакой достоверности нет в науках там, где нельзя приложить ни одну из математических наук, и в том, что не имеет связи с математикой[1].
>
> *Леонардо да Винчи*

Математика[2] является точной абстрактной наукой, изучающей специальные логические структуры, называемые математическими, у которых описаны определенные отношения между их элементами. Каждая математическая структура — аналитическая, алгебраическая, топологическая, вероятностная и другие имеют, конечно, специальные описания.

Точность математики означает, что основным методом в математических исследованиях являются логические рассуждения, а результаты исследований формулируются в строгой логической форме. Абстрактность же математики означает, что объектами ее изучения являются не материальные объекты или отношения между ними, а логические понятия и отношения между ними. Однако важно подчеркнуть, что отношения и взаимодействия между материальными объектами можно изучить с помощью их математического моделирования. Возникающие таким образом математические модели нередко приводили в свою очередь к созданию новых математических структур. Существенно то, что одна и та же математическая модель может описывать с определенным приближением свойства очень далеких друг от друга по своему конкретному содержанию реальных явлений. Для математики важна не природа рассматриваемых объектов, а лишь существующие между ними соотношения. Абстрактность математики порождает определенную трудность ее применения к описанию и решению конкретных за-

---

[1] Леонардо да Винчи. Избранные естественнонаучные произведения. М., 1955.

[2] μαθημα (греч.) — познание, наука.

дач, в то же самое время абстрактность математики придает ей силу, универсализм и общность. Роль математики, конечно, не сводится только к описанию с помощью тех или иных моделей определенных сторон каких-то явлений. Она представляет интерес и имеет большую ценность прежде всего сама по себе как наука, как знание. Математика дает мощные методы для познания мира, для изучения его закономерностей.

Математические методы исследования всегда играли и продолжают играть огромную, все увеличивающуюся роль в естествознании. В качестве примера можно привести уже ставшие хрестоматийными такие теоретические открытия, как открытие планеты Нептун, открытие электромагнитных волн или открытие позитрона, сделанные сначала математически «на кончике пера» и лишь потом нашедшие свое экспериментальное подтверждение.

Математика неустанно продолжает развиваться, в ней создаются новые методы, появляются новые разделы. Развитие математики в целом определяет уровень ее использования и оказывает существенное влияние на развитие других наук и техники. В свою очередь, задачи практики, прогресс других фундаментальных и прикладных наук приводят к созданию новых направлений математики, стимулируют ту или иную направленность математических исследований, расширяют возможность применения математических методов. В силу этого область применения математики постоянно расширяется. В последнее время благодаря появлению быстродействующих вычислительных машин в использовании математических методов произошел большой качественный скачок. Они стали применяться не только в тех областях, где математика использовалась уже давно (например, в механике, физике), но и в тех областях человеческого знания, где математика еще совсем недавно либо применялась мало, либо ее применение даже не представлялось возможным (медицина, экономика, лингвистика, социология и т. п.). Проникновение качественных и количественных математических методов в другие науки, использование в этих науках уже имеющегося математического аппарата, создание новых математических понятий и методов для описания и изучения рассматриваемых явлений, т. е. все то, что обыч-

но называется математизацией науки, является характерной чертой всего естествознания наших дней.

Современный научный работник или инженер должен в достаточной степени хорошо владеть как классическими, так и современными математическими методами исследования, которые могут применяться в его области. Для того чтобы иметь возможность с успехом применять математические методы при изучении того или иного вопроса, нужно, конечно, прежде всего иметь необходимые знания, уметь правильно обращаться с математическим аппаратом, знать границы допустимого использования рассматриваемой математической модели. Этим, однако, не исчерпываются характерные особенности решения задач математическими методами, да и вообще математического творчества, т. е. познания объективно существующих математических истин. Для правильной постановки задачи, для оценки ее данных, для выделения существенных из них и для выбора способа решения необходимо обладать еще математической интуицией, фантазией и чувством гармонии, позволяющими предвидеть нужный результат прежде, чем он будет получен. Однако интуитивно почувствовать ожидаемый результат и наметить путь исследований с помощью правдоподобных рассуждений — это далеко не все. Интуитивное чувство гармонии является в математике лишь первой, хотя и весьма важной ступенью; интуитивные соображения и правдоподобные рассуждения отдаются на суд холодного рассудка для их изучения, доказательства или опровержения. Для записи проводимых исследований и получающихся результатов используются язык цифр, разнообразные математические символы и словесные логические описания.

При математическом доказательстве гипотезы, при математическом решении задачи правильный выбор аппарата и метода — залог успеха и, более того, часто приводит к тому, что в результате получается больше полезной информации об изучаемом предмете, чем заранее предполагалось. Это связано с тем, что математический аппарат таит в себе много скрытой информации и скрытого богатства, накапливавшихся в нем в течение веков, благодаря чему формулы могут оказаться «умнее» применяющего их и дать больше, чем от них ожидалось.

Следует отметить, что в математике справедливость рассматриваемого факта доказывается не проверкой его на ряде примеров, не проведением ряда экспериментов, что не имеет для математики доказательной силы, а чисто логическим путем, по законам формальной логики.

Конечно, и эксперименты и примеры также играют большую роль в математических исследованиях: они могут или дать иллюстрацию утверждения, или опровергнуть его, или натолкнуть на какую-либо (в том числе и новую) идею. За последние годы в связи с быстрым развитием вычислительной техники особенно возросло значение математического эксперимента в прикладных исследованиях: здесь открылись качественно совершенно новые возможности и перспективы.

Безусловно, вся эта схема весьма идеализирована. Прежде всего использование знаний, математического аппарата, интуиции, чувства гармонии, фантазии, логики, эксперимента происходит не последовательно по этапам — все это все время взаимодействует между собой в течение всего процесса. Далее, далеко не всегда удается довести проводимые исследования до желаемого конца, но было бы, например, большим заблуждением думать, что для математики имеют значение только доказанные утверждения, только исследования, доведенные в известном смысле до логического завершения. Можно привести много примеров математических теорий и положений, которые, будучи сформулированы лишь в виде гипотез, тем не менее оказывали или оказывают существенное влияние на развитие математики или ее приложений.

Окончательные результаты, полученные в математике, описывая те или иные свойства логических абстрактных моделей, имеют в определенном смысле абсолютный и вечный характер и, следовательно, не меняются и не могут измениться в связи с развитием наших знаний. Так, например, за последние две тысячи лет наши представления об окружающем нас мире и об управляющих им закономерностях претерпели существенные изменения, а теорема Пифагора осталась и останется всегда такой же, какой она была в Древней Греции. Это, конечно, не исключает того, что в процессе своего исторического развития многие математические понятия и утверждения не сразу обретали и обретают свою окончательную логически законченную форму, не исключает и того, что в процессе развития одни и те же объекты изучения математи-

ки воспринимаются с разных точек зрения, что приводит к раскрытию их новых свойств, наполняет их новым содержанием, что, в свою очередь, нередко существенно меняет наше представление об их значимости и важности.

При использовании математики для описания каких-либо конкретных явлений нередко бывает достаточно лишь интуитивных представлений о соответствующих математических понятиях, однако тогда, когда математика используется в качестве метода исследования, как правило, для завершения проводимого исследования необходимо четкое представление об используемых при этом математических понятиях — только в этом случае может быть объективная уверенность в правильности сделанных выводов. Поэтому, для того чтобы применять математику как метод исследования, весьма важно осознать и хорошо освоить сущность и взаимосвязь ее основных идей и понятий, важно стремиться овладеть процессом творческого, а не формального мышления.

Свободное владение математическими методами, знания и интуиция приобретаются, накапливаются и развиваются в процессе систематических занятий, в результате длительной и настойчивой работы. Тот, кто последовательно овладевает математическим аппаратом, кто последовательно получает твердые и точные знания математических фактов, будет уверенно двигаться дальше, и математика станет послушным инструментом в его руках.

Часто мнение о трудности изучения математики связано с туманным и нечетким ее изложением на интуитивном уровне. Кажущаяся трудность тех или иных математических методов нередко обусловлена тем, что эти методы не были своевременно достаточно хорошо разъяснены и поэтому остались непонятыми.

Четкое введение математического понятия по сравнению с введением на интуитивном уровне, как правило, оправдывает себя при его применении, позволяет правильно использовать и не нуждается в дополнительных пояснениях.

Лучший и кратчайший способ в процессе обучения математике разъяснить какое-либо понятие — это дать его точную формулировку. Лучший способ на первом этапе обучения объяснить теорему, выяснить ее смысл, установить ее связь с ранее изученными фактами — это доказать теорему. Сделать это надо просто, естественно и доходчиво, что часто

совсем не легко. В умении осуществить это на достаточно высоком уровне и состоит прежде всего и с к у с с т в о преподавания математики. Однако было бы неправильно думать, что с овладением доказательством математической теоремы кончается процесс ее познания. До конца смысл и роль теорем раскрываются лишь при их применении к изучению других теоретических вопросов и решении тех или иных конкретных задач. Трудно переоценить огромную роль анализа отдельных примеров, иллюстрирующих теоретические утверждения, и решения с помощью последних соответствующих частных задач. Безусловно, при достаточно хорошей математической культуре вполне допустимо знакомство с рядом утверждений, ограничивающееся лишь их формулировкой без проведения доказательства. Однако на первом этапе обучения это явно нецелесообразно.

Косвенная польза от изучения математики состоит в том, что оно совершенствует общую культуру мышления, дисциплинирует ее, приучает человека логически рассуждать, воспитывает у него точность и обстоятельность аргументации. Математика учит не загромождать исследование ненужными подробностями, не влияющими на сущность дела, и, наоборот, не пренебрегать тем, что имеет принципиальное значение для существа изучаемого вопроса. Все это дает возможность эффективно исследовать и осмысливать новые задачи, возникающие в различных областях человеческой деятельности.

Умение логически мыслить, владение математическим аппаратом, правильное использование математики дают большую экономию мышления, вооружают человека мощным методом исследования.

Овладеть в достаточной мере математическим методом, математической культурой мышления, почувствовать силу и красоту математических методов — далеко не простая задача. Но для того, кто сумеет этого достичь, труд не пропадет зря. Для него откроются новые перспективы человеческой деятельности, заманчивые дороги в неизвестное, откроются качественно новые возможности творчества и познания мира. Причем важно отметить, что все это доступно каждому, кто хочет овладеть математикой, кто серьезно и последовательно займется ее изучением.

Глава 1

Дифференциальное исчисление функций одной переменной

§ 1.

Множества и функции. Логические символы

1.1. Множества. Операции над множествами

В математике первичными понятиями являются понятия множества, элемента и принадлежности элемента множеству. Множества далее будем обозначать большими буквами латинского или какого-либо другого алфавита: A, B, ... , X, Y, ... , $\mathfrak{B}$, $\mathfrak{B}$, ... , а элементы множеств — малыми буквами: a, b, ... , x, y, ... , α, β, ... Если a является элементом множества A, то пишут $a \in A$ (читается: «a принадлежит множеству A») или, что то же, $A \ni a$. Если же a не принадлежит множеству A, то пишут $a \notin A$ или $A \not\ni a$.

Множества A и B называются *равными*, если они состоят из одних и тех же элементов. Таким образом, равенство $A = B$ означает применительно к множествам, что одно и то же множество обозначено разными буквами A и B.

Запись $A = \{a, b, c, ...\}$ означает, что множество A состоит из элементов a, b, c и, возможно, каких-то других, заданных тем или иным способом.

Если множество A состоит из элементов a_α, где α пробегает некоторое множество индексов $\mathfrak{A}$, то будем писать $A = \{a_\alpha\}$ или, подробнее, $A = \{a_\alpha\}$, $\alpha \in \mathfrak{A}$, или, если это не может привести к недоразумению, просто $A = \{a\}$. Если множество A состоит из всех элементов, обладающих определенным свойством, то будем писать $A = \{a: ...\}$, где в фигурных скобках после двоеточия записано указанное свойство элементов множества A. Например, если a и b — два таких действи-

тельных числа, что $a \leqslant b$, и через $[a, b]$ обозначено множество всех действительных чисел x, удовлетворяющих неравенствам $a \leqslant x \leqslant b$, то определение этого множества (отрезка) с помощью введенных символов можно записать следующим образом:

$$[a, b] = \{x\colon a \leqslant x \leqslant b\}.$$

Для удобства вводится понятие множества, не содержащего никаких элементов. Оно называется *пустым множеством* и обозначается символом $\varnothing$. По определению, оно не содержит элементов, но причисляется к множествам.

Если каждый элемент множества A является элементом множества B, то говорят, что множество A есть часть множества B или что A является *подмножеством* множества B, и пишут $A \subset B$ (читается: «множество A содержится во множестве B») или, что то же, $B \supset A$ (читается: «множество B содержит множество A»).

УПРАЖНЕНИЕ 1. Доказать, что включения $A \subset B$ и $B \subset A$ выполняются одновременно тогда и только тогда, когда $A = B$.

Из определения подмножества следует, что $A \subset A$, каково бы ни было множество A; принято также считать, по определению, что пустое множество является подмножеством каждого множества: $\varnothing \subset A$. Если A — произвольное множество, то $\varnothing$ и A называются его *несобственными подмножествами*; если же $A \subset B$ и существует такой элемент $x \in B$, что $x \notin A$, то множество A называется *собственным подмножеством* множества B.

Если заданы два множества A и B (рис. 1, *а*), то через $A \cup B$ обозначается множество, называемое их *объединением* или *суммой* и состоящее из всех тех элементов, каждый из которых принадлежит хотя бы одному из множеств A и B (рис. 1, *б*). Таким образом, если некоторый элемент принадлежит множеству $A \cup B$, то он принадлежит либо только

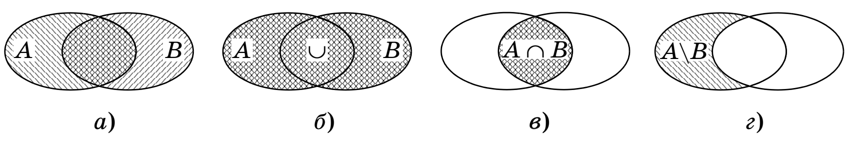


Рис. 1

множеству A, либо только множеству B, либо обоим этим множествам.

Для любого множества A (непустого или пустого) полагается $A \cup \varnothing = A$.

Через $A \cap B$ обозначается множество, состоящее из всех элементов, которые одновременно принадлежат как множеству A, так и множеству B (рис. 1, *в*). Множество $A \cap B$ называется *пересечением множеств* A и B. Если A и B не имеют общих элементов (в частности, одно из них или оба пусты), то полагают $A \cap B = \varnothing$. В этом случае множества A и B называются *непересекающимися*.

Отметим, что пустое множество совпадает само с собой: $\varnothing = \varnothing$, но вместе с тем оно не пересекается само с собой: $\varnothing \cap \varnothing = \varnothing$.

Через $A \setminus B$ обозначается множество, называемое *разностью множеств* A и B и состоящее из всех элементов, которые принадлежат множеству A, но не принадлежат множеству B (рис. 1, *г*). Говорят также, что $A \setminus B$ получается из множества A *вычитанием* из него множества B.

Если $B \subset A$, то разность $A \setminus B$ называется *дополнением множества B до множества* A или дополнением B в A. По определению, полагается $A \setminus A = \varnothing$.

УПРАЖНЕНИЕ 2. Доказать, что $A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$.

Если задана система множеств A_α (термины «множество», «система», «совокупность», «класс» будут употребляться как синонимы), где значения α образуют некоторую совокупность индексов $\mathfrak{A}$, то *объединением* $\bigcup\limits_{\alpha \in \mathfrak{A}} A_\alpha$ *множеств* A_α называется множество, каждый элемент которого принадлежит хотя бы одному из заданных множеств A_α, т. е. условие $x \in \bigcup\limits_{\alpha \in \mathfrak{A}} A_\alpha$ равносильно следующему: существует такое $\alpha \in \mathfrak{A}$, что $x \in A_\alpha$.

Пересечением множеств A_α, $\alpha \in \mathfrak{A}$, называется такое множество, обозначаемое через $\bigcup\limits_{\alpha \in \mathfrak{A}} A_\alpha$, что каждый его элемент принадлежит всем множествам A_α, т. е. условие $x \in \bigcup\limits_{\alpha \in \mathfrak{A}} A_\alpha$ означает: для всех $\alpha \in \mathfrak{A}$ имеет место $x \in A_\alpha$.

Для любой системы множеств $\{A_\alpha\}$, $A_\alpha \subset X$, $\alpha \in \mathfrak{A}$, и любого множества X справедливы следующие соотношения:

$$X \setminus \bigcup_{\alpha \in \mathfrak{A}} A_\alpha = \bigcap_{\alpha \in \mathfrak{A}} (X \setminus A_\alpha); \tag{1.1}$$

$$X \setminus \bigcap_{\alpha \in \mathfrak{A}} A_\alpha = \bigcup_{\alpha \in \mathfrak{A}} (X \setminus A_\alpha). \tag{1.2}$$

Докажем, например, равенство (1.1). Если $x \in X \setminus \bigcup_{\alpha \in \mathfrak{A}} A_\alpha$, то, в силу определения разности множеств, $x \in X$ и $x \notin \bigcup_{\alpha \in \mathfrak{A}} A_\alpha$. В свою очередь, это, согласно определению объединения множеств, эквивалентно тому, что $x \in X$ и для всех $\alpha \in \mathfrak{A}$ имеет место соотношение $x \notin A_\alpha$. Это же, снова по определению разности множеств, равносильно утверждению, что для всех $\alpha \in \mathfrak{A}$ имеем $x \in X \setminus A_\alpha$. Наконец, последнее утверждение, по определению пересечения множеств, означает, что $x \in \bigcap_{\alpha \in \mathfrak{A}} (X \setminus A_\alpha)$. Итак, условия $x \in X \setminus \bigcup_{\alpha \in \mathfrak{A}} A_\alpha$ и $x \in \bigcap_{\alpha \in \mathfrak{A}} (X \setminus A_\alpha)$ эквивалентны, поэтому выполняется равенство (1.1). Равенство (1.2) доказывается аналогично. Подобными же рассуждениями доказывается справедливость и следующих равенств для любых множеств A, B, C:

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C),$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C).$$

В пункте 1.2* рассмотрено понятие функции, а пункт 1.3* будет посвящен понятиям конечного множества и последовательности. Пункты и параграфы курса, отмеченные звездочкой, при первом чтении можно опустить и вернуться к ним лишь в случае внутренней потребности. В частности, для понимания дальнейшего материала достаточно имеющегося в курсе элементарной математики представления о функции как об определенном соответствии между элементами двух множеств. При этом понятие соответствия следует понимать как первичное.

1.2*. Функции

Будем говорить, что *число элементов множества A равно единице*, если в нем имеется элемент $a \in A$ и нет других (ина-

че говоря, если из множества A вычесть множество, состоящее из элемента a, то получится пустое множество).

Множество A называется *множеством из* 2 (*двух*) *элементов*, если после вычитания из него множества, состоящего только из одного элемента $a \in A$, т. е. множества, число элементов которого равно 1, останется множество, число элементов которого также равно единице. Нетрудно убедиться, что это определение не зависит от выбора указанного элемента $a \in A$, т. к. если $a \in A$ и $b \in A$, причем $A \setminus \{a\}$ состоит из одного элемента b, то и множество $A \setminus \{b\}$ также состоит из одного элемента (а именно из элемента a).

Пусть теперь заданы множества $X = \{x\}$ и $Y = \{y\}$. Множество, состоящее из двух элементов $x \in X$ и $y \in Y$, называется *парой* $\{x, y\}$ элементов x, y.

Пара вида $\{x, \{x, y\}\}$, где $x \in X$, $y \in Y$, $\{x, y\}$ — пара элементов x, y, называется *упорядоченной парой* элементов x и y. Элемент x называется первым элементом упорядоченной пары $\{x, \{x, y\}\}$, а элемент y — вторым. Упорядоченная пара $\{x, \{x, y\}\}$ обозначается через (x, y). В дальнейшем под парой обычно понимается упорядоченная пара.

Множество всех упорядоченных пар (x, y), $x \in X$, $y \in Y$, называется *произведением множеств* X и Y и обозначается через $X \times Y$. При этом не предполагается, что обязательно множество X отлично от множества Y, т. е. возможен и случай, когда $X = Y$.

Определение 1. *Всякое множество* $f = \{(x, y)\}$ *упорядоченных пар* (x, y), $x \in X$, $y \in Y$ *такое, что для любых пар* $(x', y') \in f$ *и* $(x'', y'') \in f$ *из условия* $y' \neq y''$ *следует, что* $x' \neq x''$, *называется функцией или, что то же, отображением.*

Наряду с терминами «функция» и «отображение» в определенных ситуациях употребляются им равнозначные термины «преобразование», «морфизм», «соответствие». Функции будем обозначать различными буквами: f, g, ... , F, G, ... , φ, ψ, ...

Множество всех первых элементов упорядоченных пар (x, y) некоторой функции f называют *множеством задания* или *множеством* (*областью*) *определения* этой функции и обозначают через X_f, а множество всех вторых элементов — *множеством ее значений* и обозначают через Y_f.

Очевидно, что $X_f \subset X$, $Y_f \subset Y$. Само множество упорядоченных пар $f = \{(x, y)\}$, рассматриваемое как подмножество произведения $X \times Y$, называется *графиком функции* f.

Элемент $x \in X_f$ называется *аргументом функции* f или *независимой переменной*, а элемент $y \in Y$ — *зависимой переменной*.

Если $f = \{(x, y)\}$ есть функция (отображение), то пишут $f : X_f \to Y$ и говорят, что f отображает множество X_f в множество Y. В случае $X = X_f$ пишется просто $f : X \to Y$.

Если $f : X \to Y$ — функция, т. е. множество упорядоченных пар $f = \{(x, y)\}$, $x \in X$, $y \in Y$, удовлетворяющих условиям определения 1, и $(x, y) \in f$, то пишут $y = f(x)$ (иногда просто $y = fx$) или $f : x \mapsto y$ и говорят, что функция f ставит в соответствие элементу x элемент y (отображение f отображает элемент x в элемент y) или, что то же самое, элемент y соответствует элементу x.

Для пары (x, y), где $y = f(x)$, говорят также, что элемент y является значением функции f в точке x или образом элемента x при отображении f.

Наряду с символом $f(x_0)$ для обозначения значения функции f в точке x_0 употребляют также обозначение $f(x)|_{x = x_0}$.

Иногда саму функцию f обозначают символом $f(x)$. Обозначение функции $f : X \to Y$ и ее значения в точке $x \in X$ одним и тем же символом $f(x)$ не приводит к недоразумению, так как в каждом конкретном случае всегда ясно, о чем именно идет речь.

Обозначение $f(x)$ обычно удобнее обозначения $f : x \mapsto y$ при вычислениях. Например, запись $f(x) = x^2$ значительно удобнее и проще использовать при аналитических преобразованиях, чем запись $f : x \mapsto x^2$.

При заданном $y \in Y$ совокупность всех таких элементов $x \in X$, что $f(x) = y$, называют *прообразом элемента* y и обозначают $f^{-1}(y)$. Таким образом, $f^{-1}(y) = \{x : x \in X, f(x) = y\}$.

Очевидно, если $y \in Y \setminus Y_f$, то $f^{-1}(y) = \varnothing$.

Пусть задано отображение $f : X \to Y$, т. е. отображение множества X в множество Y. Иначе говоря, каждому элементу $x \in X$ поставлен в соответствие, и притом единствен-

ный, элемент $y \in Y$ и каждый элемент $y \in Y_f \subset Y$ поставлен в соответствие хотя бы одному элементу $x \in X$.

Если $Y = X$, то говорят, что отображение f отображает множество X в себя.

Если $Y = Y_f$, т. е. множество Y совпадает с множеством значений функции f, то говорят, что f отображает множество X *на* множество Y или что отображение f является сюръективным отображением, короче — *сюръекцией*. Таким образом, отображение $f : X \to Y$ есть сюръекция, если для любого элемента $y \in Y$ существует по крайней мере один такой элемент $x \in X$, что $f(x) = y$.

Очевидно, если $f : X \to Y$ и Y_f — множество значений функции f, то $f : X \to Y_f$ является сюръективным отображением.

Если при отображении $f : X \to Y$ разным $x \in X$ соответствуют разные $y \in Y$, т. е. при $x' \neq x''$ имеет место $f(x') \neq f(x'')$, то отображение f называется взаимно однозначным отображением (взаимно однозначным соответствием) X в Y, а также *однолистным отображением* или *инъекцией*. Таким образом, отображение $f : X \to Y$ однолистно (инъективно) тогда и только тогда, когда прообраз каждого элемента y, принадлежащего множеству значений функции $f : y \in Y_f$, состоит в точности из одного элемента.

Если отображение $f : X \to Y$ является одновременно взаимно однозначным и на множество Y, т. е. является одновременно инъекцией и сюръекцией, то оно естественно называется *взаимно однозначным отображением* множества X *на* множество Y или, что то же, *биективным отображением* (биекцией) в Y.

Таким образом, отображение $f : X \to Y$ является взаимно однозначным отображением множества X на множество Y тогда и только тогда, когда для любых $x' \in X$ и $x'' \in X$, $x' \neq x''$, справедливо неравенство $f(x') \neq f(x'')$, и, каково бы ни было $y \in Y$, существует такой элемент $x \in X$, что $f(x) = y$.

Взаимно однозначное отображение множества X на множество Y часто называют также взаимно однозначным соответствием элементов этих множеств.

Если $f : X \to Y$ и $A \subset X$, то множество

$$S = \{y : y \in Y,\ y = f(x),\ x \in A\},$$

т. е. множество всех тех y, в каждый из которых при отображении f отображается хотя бы один элемент из подмножества A множества X, называют *образом подмножества* A и пишут $S = f(A)$. В частности, всегда имеем $Y_f = f(X)$.

Для образов множеств $A \subset X$ и $B \subset X$ справедливы следующие легко проверяемые соотношения:

$$f(A \cup B) = f(A) \cup f(B),$$
$$f(A \cap B) \subset f(A) \cap f(B),$$
$$f(A) \setminus f(B) \subset f(A \setminus B),$$

а если $A \subset B$, то $f(A) \subset f(B)$.

Если $f : X \to Y$ и $B \subset Y$, то множество

$$A = \{x : x \in X,\ f(x) \in B\}$$

называют *прообразом множества* B и пишут $A = f^{-1}(B)$. Таким образом, прообраз множества B состоит из всех тех элементов $x \in X$, которые при отображении f отображаются в элементы из B или, что то же самое, он состоит из всех прообразов точек $y \in B$:

$$f^{-1}(B) = \bigcup_{y \in B} f^{-1}(y).$$

Для прообразов множеств $A \subset Y$ и $B \subset Y$ справедливы следующие также легко доказываемые соотношения:

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B),$$
$$f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B),$$
$$f^{-1}(A \setminus B) = f^{-1}(A) \setminus f^{-1}(B),$$

а если $A \subset B$, то $f^{-1}(A) \subset f^{-1}(B)$.

Если $A \subset X$, то функция $f : X \to Y$ естественным образом порождает функцию, определенную на множестве A, ставящую в соответствие каждому элементу $x \in A$ элемент $f(x)$. Эту функцию называют *сужением функции* f на множество A и иногда обозначают через $f|_A$ или просто через f_A. Таким образом, f_A отображает A в Y и для любого $x \in A$ имеет место $f_A : x \mapsto f(x)$. Если множество A не совпадает с множеством X, т. е. является собственным подмножеством множества X, то сужение f_A функции f на множество A имеет другую область определения, чем функция f, и, следовательно, является другой, чем f, функцией. Нередко сужение функции $f : X \to$

$\to Y$ на множество $A \subset X$ обозначают тем же символом f и называют «функцией f на множестве A».

Если две функции f и g рассматриваются на одном и том же множестве X, точнее, если рассматриваются сужения функций f и g на одно и то же множество X, то запись $f \equiv g$ на X означает, что $f(x) = g(x)$ для каждого $x \in X$. В этом случае говорят, что функция f тождественно равна функции g на множестве X.

Иногда знак «$\equiv$» будет употребляться между символами, обозначающими один и тот же объект (обычно один из употребленных в этом случае символов содержит более подробное описание объекта, чем другой символ), например запись $f(x) \equiv f(x_1, \ldots, x_n)$ означает, что $x = (x_1, \ldots, x_n)$, и тем самым запись $f(x)$ и $f(x_1, \ldots, x_n)$ обозначает одну и ту же функцию.

Отметим, что функции, у которых всем элементам некоторого множества соответствует один и тот же элемент, т. е. функции, у которых при изменении значения аргумента значение функции не меняется, называются *постоянными* (на данном множестве) или *константами*.

Итак, если при изменении одной переменной (аргумента функции) другая переменная, являющаяся функцией первой, не меняется (т. е. «не зависит» от первой переменной), то это частный и в определенном смысле простейший случай функциональной зависимости.

Если $f : X \to Y$ и каждый элемент $y \in Y_f$ представляет собой множество каких-то элементов $y = \{z\}$, причем среди этих множеств $\{z\}$ имеется по крайней мере одно непустое множество, состоящее не из одного элемента, то такая функция f называется *многозначной функцией*. При этом элементы z множества $f(x) = \{z\}$ часто также называют значениями функции f в точке x.

Если каждое множество $f(x)$ состоит только из одного элемента, то функцию f называют также *однозначной функцией*.

Если $f : X \to Y$ и $g : Y \to Z$, то функция $F : X \to Z$, определенная для каждого $x \in X$ равенством $F(x) = g(f(x))$, называется *композицией* (иногда *суперпозицией*) функций f и g или *сложной функцией* и обозначается через $g \circ f$.

Таким образом, по определению для каждого $x \in X$

$$(g \circ f)(x) = g(f(x)).$$

Пусть задана функция $f : X \to Y$ и Y_f — множество ее значений. Совокупность всевозможных упорядоченных пар вида $(y, f^{-1}(y))$, $y \in Y_f$, образует функцию, которая называется *обратной функцией* для функции f и обозначается через f^{-1}. Обратная функция f^{-1} ставит в соответствие каждому элементу $y \in Y_f$ его прообраз $f^{-1}(y)$, т. е. некоторое множество элементов. Тем самым обратная функция является, вообще говоря, многозначной функцией. При этом для любой точки $x \in X$ имеет место включение $f^{-1}(f(x)) \ni x$, а для любой точки $y \in Y_f$ — равенство $f(f^{-1}(y)) = y$.

Отображения f и f^{-1} называются *взаимно обратными*.

Если отображение $f : X \to Y$ однолистно (инъективно), то обратное отображение, определенное, как всегда на Y_f, является однозначной функцией и отображает Y_f на X, т. е. $f^{-1} : Y_f \to X$. Действительно, в этом случае прообразы всех точек $y \in Y_f$ состоят точно из одной точки $x \in X$.

Очевидно, что для инъективного отображения обратное отображение f^{-1} удовлетворяет следующим условиям: при любом $x \in X$ имеет место равенство $f^{-1}(f(x)) = x$, а при любом $y \in Y_f$ — равенство $f(f^{-1}(y)) = y$. При этом для заданного инъективного отображения f каждое из указанных двух условий однозначно определяет обратное отображение f^{-1}.

1.3*. Конечные множества и натуральные числа. Последовательности

В определенном смысле простейшими множествами являются так называемые конечные множества. Определим это понятие.

В п. 1.2* были определены понятия множеств, состоящих из одного и двух элементов. Напомним эти определения, снабдив их некоторыми комментариями. Если множество X содержит некоторый элемент x и не содержит никакого другого, т. е. после вычитания из множества X множества, состоящего только из элемента x, получается пустое множество

$$X \setminus \{x\} = \varnothing,$$

то множество X называется множеством, состоящим из одного элемента. Таким образом, понятие множества, состоящего из одного элемента, равносильно понятию элемента. Очевидно, что два множества, состоящие из одного элемента, можно отобразить друг на друга взаимно однозначно.

Говорят, что число элементов каждого множества, состоящего из одного элемента, равно единице, и слово «единица» обозначают символом 1.

Если множество X после вычитания из него множества, состоящего из одного какого-либо элемента $x \in X$, превращается в множество, состоящее только из одного элемента $y \in X$, т. е. $X \setminus \{x\} = \{y\}$, то множество X называется множеством, состоящим из двух элементов. Ясно, что любые два множества, состоящие из двух элементов, также можно отобразить друг на друга взаимно однозначно. Говорят, что число элементов каждого множества, состоящего из двух элементов, равно двум, и слово «два» обозначают символом 2.

Если множество X после вычитания из него множества, состоящего из одного элемента $x \in X$, превращается в множество, состоящее из двух элементов, то множество X называют множеством, состоящим из трех элементов, или множеством, число элементов которого равно трем, и слово «три» обозначают символом 3.

Аналогично, последовательно определяются множества, состоящие из четырех, пяти, шести и т. д. элементов, или, что то же самое, множества, число элементов которых равно четырем, пяти, шести и т. д. Слова «четыре», «пять», «шесть» и т. д. обозначаются символами 4, 5, 6 и т. д.

Элементы множества

$$\{1, 2, 3, 4, 5, 6, \ldots\} \tag{1.3}$$

называются *натуральными числами*. Множество всех натуральных чисел обозначают через $\boldsymbol{N}$.

Обозначим через n произвольно фиксированное натуральное число: $n \in \boldsymbol{N}$, а натуральное число, непосредственно следующее за числом n в множестве (1.3), обозначим через $n + 1$. Число n будем называть предшествующим числу $n + 1$. Очевидно, что всякое натуральное число n, кроме 1, имеет предшествующее, которое будем обозначать $n - 1$. Таким образом, $(n - 1) + 1 = n$.

Согласно определению натуральных чисел, каждое натуральное число n представимо в виде

$$n = (\ldots((1 + 1) + 1 \ldots + 1) + 1. \tag{1.4}$$

(Рассматривая множества, состоящие из «одинаковых» элементов, можно сказать, что в формуле в правой части стоит множество из n единиц, соединенных последовательно знаком «+».)

Согласно конструкции, определяющей натуральные числа, число $n + 1 \in \boldsymbol{N}$ характеризуется тем свойством, что каждое множество, состоящее из $n + 1$ элемента, после удаления любого из них превращается в множество, которое состоит из n элементов.

Согласно той же конструкции, любые два множества, состоящие из $n \in \boldsymbol{N}$ элементов, отображаются друг на друга взаимно однозначно.

Если из двух натуральных чисел m и n число m встречается раньше, чем n в ряде натуральных чисел (1.3), т. е. число m стоит левее числа n, то число m называют *меньшим числа* n и пишут $m < n$ или, что то же, число n называют *бóльшим числа* m и пишут $n > m$.

Например, $n - 1 < n < n + 1$, $n \neq 1$. Для любых двух различных натуральных чисел m и n имеет место точно одно из соотношений $m < n$ или $m > n$. При этом если $m < n$ и $n < p$, то $m < p$, $m, n, p \in \boldsymbol{N}$.

Если число $m \in \boldsymbol{N}$ меньше числа $n \in \boldsymbol{N}$, то в каждом множестве, состоящем из n элементов, имеются подмножества, состоящие из m элементов. Это следует из самой конструкции последовательного определения натуральных чисел.

Множество натуральных чисел $\boldsymbol{N}$ обладает следующим замечательным свойством.

Если множество $\boldsymbol{M}$ *таково, что*:

1) $\boldsymbol{M} \subset \boldsymbol{N}$;

2) $1 \in \boldsymbol{M}$;

3) *из* $n \in \boldsymbol{M}$ *следует, что* $n + 1 \in \boldsymbol{M}$, *то*

$$\boldsymbol{M} = \boldsymbol{N}. \tag{1.5}$$

Действительно, согласно условию 2), имеем $1 \in \boldsymbol{M}$, поэтому, согласно свойству 3), и $2 \in \boldsymbol{M}$; тогда, согласно тому же свойству 3), получаем $3 \in \boldsymbol{M}$. Но любое натуральное число $n \in \boldsymbol{N}$ получается из 1 последовательным прибавлением к ней той

же 1, поэтому $n \in \boldsymbol{M}$, т. е. $\boldsymbol{N} \subset \boldsymbol{M}$, а так как по условию 1) выполняется включение $\boldsymbol{M} \subset \boldsymbol{N}$, то $\boldsymbol{M} = \boldsymbol{N}$. Итак, равенство (1.5) доказано.

Из всего сказанного следует, что множество натуральных чисел $\boldsymbol{N}$ обладает следующими свойствами.

1^0. *Каждому элементу* $n \in \boldsymbol{N}$ *поставлен в соответствие точно один элемент этого множества, обозначаемый через* $n + 1$ *и называемый элементом, следующим за элементом* n.

2^0. *Каждый элемент из* $\boldsymbol{N}$ *может следовать только за одним элементом* $n \in \boldsymbol{N}$.

3^0. *Существует единственный элемент, обозначаемый символом* 1, *который не следует ни за каким элементом.*

4^0. *Если множество* $\boldsymbol{M} \subset \boldsymbol{N}$ *таково, что* $1 \in \boldsymbol{M}$, *и из включения* $m \in \boldsymbol{M}$ *следует, что* $m + 1 \in \boldsymbol{M}$, *то* $\boldsymbol{M} = \boldsymbol{N}$.

Справедливо и обратное утверждение в том смысле, что всякое множество, удовлетворяющее условиям 1^0—4^0, может быть взаимно однозначно отображено на множество $\boldsymbol{N}$ с сохранением соотношения «больше — меньше».

Это обстоятельство позволяет получить аксиоматическое определение множества натуральных чисел: достаточно принять утверждения 1^0—4^0 за аксиомы (они называются аксиомами Пеано[1]).

Таким образом, аксиоматическое определение множества натуральных чисел выглядит следующим образом.

Определение 2. *Множество, удовлетворяющее условиям* 1^0—4^0, *называется множеством натуральных чисел.*

При таком определении свойство 4^0 называется *аксиомой индукции*.

Из доказанного равенства (1.5) следует так называемый принцип доказательства *методом математической индукции*.

Если имеется множество утверждений, каждому из которых приписано натуральное число (его номер) $n = 1, 2, \ldots$, *и если доказано, что*:

1) *справедливо утверждение с номером* 1;

---

[1] Д. Пеано (1858—1932) — итальянский математик.

2) *из справедливости утверждения с любым номером $n \in \boldsymbol{N}$ следует справедливость утверждения с номером $n + 1$, то тем самым доказана справедливость всех рассматриваемых утверждений, т. е. справедливость утверждения с произвольным номером $n \in \boldsymbol{N}$.*

Пример доказательства, проведенного методом математической индукции, см., например, в п. 1.4 теорему 1.

Про пустое множество говорят, что число его элементов равно нулю. Слово «нуль» обозначается символом 0. Множество натуральных чисел, к которому добавлен нуль, обозначается $\boldsymbol{N}_0$:

$$\boldsymbol{N}_0 \overset{\text{def}}{=} \boldsymbol{N} \cup \{0\}. \qquad (1.6)$$

Нуль считается меньшим любого натурального числа: $0 < n$, $n \in \boldsymbol{N}$.

Натуральные числа можно складывать между собой и с нулем: *операция сложения* в множестве $\boldsymbol{N}_0$ определяется с помощью операции объединения множеств. *Суммой* $m + n$ чисел $m \in \boldsymbol{N}$ и $n \in \boldsymbol{N}$ называется число элементов объединения двух непересекающихся множеств, одно из которых содержит m, а другое — n элементов.

Это определение не зависит от выбора указанных множеств. Легко видеть, что в случае $n = 1$ число $m + n = m + 1$ в смысле операции сложения совпадает со следующим за m числом, обозначенным раньше тем же символом $m + 1$.

Также непосредственно из определения вытекают следующие свойства операции сложения в множестве $\boldsymbol{N}_0$.

1^0. $n + 0 = n$, $n \in \boldsymbol{N}_0$.

2^0. З а к о н к о м м у т а т и в н о с т и:

$$m + n = n + m, \; m, \; n \in \boldsymbol{N}_0.$$

3^0. З а к о н а с с о ц и а т и в н о с т и:

$$m + (n + p) = (m + n) + p, \; m, \; n, \; p \in \boldsymbol{N}_0.$$

Операция вычитания в множестве $\boldsymbol{N}_0$ определяется с помощью операции вычитания множеств.

Если $m \leqslant n$[1], $m, \; n \in \boldsymbol{N}_0$, то разностью $m - n$ называется число элементов множества, которое получается из множе-

---

[1] Т. е. либо $m = n$, либо $m < n$.

ства, содержащего n элементов, с помощью вычитания из него множества, содержащего m элементов. Это определение не зависит от выбора рассматриваемых множеств, содержащих n и m элементов.

Операция вычитания в множестве $\boldsymbol{N}_0$ является обратной для операции сложения: если

$$m + n = p,\ m,\ n \in \boldsymbol{N}_0,$$

то

$$n = p - m.$$

Из определения натурального числа n следует, что множество $\{1, 2, \ldots, n\}$ состоит из n элементов. Поэтому каждое множество, состоящее из n элементов, взаимно однозначно отображается на множество $\{1, 2, \ldots, n\}$.

Определение 3. *Если для множества существует такое натуральное n, что число его элементов равно n, то такое множество называется конечным* (*состоящим из n элементов*).

Всякое множество, не являющееся конечным, называется *бесконечным*. Примером бесконечного множества является множество $\boldsymbol{N}$ всех натуральных чисел.

Пустое множество, по определению, причисляется к конечным множествам.

УПРАЖНЕНИЕ 3. Доказать равенство

$$m + n = \underbrace{(\ldots((m + 1) + 1)\ldots + 1) + 1}_{n \text{ раз}}. \qquad (1.7)$$

Выражение «n раз» означает здесь, что берется множество из n элементов, каждый из которых обозначен 1, и все эти единицы последовательно прибавляются к натуральному числу m.

Определение 4. *Пусть X — какое-либо множество и $\boldsymbol{N}$ — множество натуральных чисел. Всякое отображение $f : \boldsymbol{N} \to X$* (см. п. 1.2*) *называется последовательностью элементов множества X. Элемент $f(n)$ обозначается через x_n и называется n-м членом последовательности $f : \boldsymbol{N} \to X$, а сама эта последовательность обозначается через $\{x_n\}$ или x_n, $n = 1, 2, \ldots$.*

Каждый элемент x_n последовательности $\{x_n\}$ представляет собой упорядоченную пару, состоящую из числа $n \in \boldsymbol{N}$ и соответствующего ему при отображении $f : \boldsymbol{N} \to X$ элемента x множества X, т. е. $x_n = (n, x)$. Второй элемент этой пары называется *значением элемента* x_n последовательности $\{x_n\}$, а первый — его *номером*.

Множество элементов последовательности всегда бесконечно. Два различных элемента последовательности могут иметь одно и то же значение, но заведомо отличаются номерами, которых бесконечное множество.

Множество значений элементов последовательности (обычно говорят короче: множество значений последовательности) может быть конечным. Например, если всем $n \in \boldsymbol{N}$ поставлен в соответствие один и тот же элемент $a \in X$, т. е. при всех $n \in \boldsymbol{N}$ имеет место равенство $f(n) = a$, то множество значений последовательности $x_n = a$, $n = 1, 2, \ldots$, состоит из одного элемента $a \in X$. Такие последовательности называются *стационарными*.

Если $n_1 < n_2$, $n_1 \in \boldsymbol{N}$, $n_2 \in \boldsymbol{N}$, то член x_{n_1} последовательности $\{x_n\}$ называется членом, предшествующим члену x_{n_2}, а член x_{n_2} — членом, следующим за членом x_{n_1}. В этом смысле члены последовательности всегда упорядочены.

Иногда в качестве номеров бывает удобно употреблять не все натуральные числа, а лишь некоторые из них. Например, натуральные числа начиная с некоторого натурального числа n_0: x_n, $n = n_0, n_0 + 1 \ldots$, или одни четные числа: x_n, $n = 2, 4, \ldots$. Случается, что для нумерации употребляются не только натуральные, но и другие числа, например x_n, $n = 0, 1, 2, \ldots$ (здесь в качестве еще одного номера добавлен нуль). Во всех этих случаях можно перенумеровать заново x_n, используя все натуральные числа m и только их. В первом примере следует положить $m = n - n_0 + 1$, во втором — $m = \frac{n}{2}$, в третьем — $m = n + 1$. Поэтому в подобных случаях также говорят, что x_n образуют последовательность, и, конечно, указывают, какие значения принимают номера n.

1.4. Группировки элементов конечного множества

Если задано какое-либо конечное множество, то бывает полезным составить из его элементов некоторые группы, удовлетворяющие тем или иным условиям, и изучить их свойства, например выяснить, сколько всего можно составить таких групп. Рассмотрим группы элементов, называемые размещениями, перестановками и сочетаниями.

Пусть задано множество, состоящее из n элементов,

$$x_1, x_2, \dots, x_n \tag{1.8}$$

и фиксировано некоторое натуральное число $k \leqslant n$. В дальнейшем в настоящем пункте под элементами всегда будут пониматься элементы (1.8), если, конечно, не оговорено что-либо другое.

Определение 5. *Группы элементов, состоящие из k элементов в каждой и отличающиеся друг от друга либо самими элементами, либо их порядком, называются размещениями из n элементов по k.*

Например, группы $\{1, 2\}, \{2, 1\}, \{1, 3\}, \{3, 1\}, \{2, 3\}, \{3, 2\}$ образуют всевозможные размещения из трех натуральных чисел 1, 2, 3 по два из них.

Число всех размещений из n элементов по k элементов обозначается через A_n^k, $n = 1, 2, \dots, k = 1, 2, \dots, n$.

ЛЕММА. *Если $k < n$, то*

$$A_n^{k+1} = A_n^k (n - k). \tag{1.9}$$

Доказательство. Из каждого размещения по k элементов из n, т. е. из каждой упорядоченной группы k элементов $\{x_{i_1}, x_{i_2}, \dots, x_{i_k}\}$, добавлением одного из не входящих в него элемента $x_{i_{k+1}}$ (таких элементов всего $n - k$) можно получить $n - k$ размещений по $k + 1$ элементов вида $\{x_{i_1}, x_{i_2}, \dots$ $\dots, x_{i_k}, x_{i_{k+1}}\}$, причем таким способом получаются все размещения по $k + 1$ элементов и точно по одному разу. Поэтому

$$A_n^k (n - k) = A_n^{(k+1)}.$$

ТЕОРЕМА 1. *Имеет место формула*

$$A_n^k = n(n-1)\,\dots\,(n-k+1),\ k = 1, 2, \dots, n. \qquad (1.10)$$

Доказательство. Очевидно, что

$$A_n^1 = n. \qquad (1.11)$$

Далее,

$$A_n^2 \underset{(1.9)}{=} A_n^{1+1} = A_n^1(n-1) \underset{(1.11)}{=} n(n-1), \qquad (1.12)$$

$$A_n^3 \underset{(1.9)}{=} A_n^{2+1} = A_n^2(n-2) \underset{(1.12)}{=} n(n-1)(n-2).$$

Вообще, если

$$A_n^{(k-1)} = n(n-1)\,\dots\,(n-k+2), \qquad (1.13)$$

то

$$A_n^k \underset{(1.9)}{=} A_n^{k-1}(n-(k-1)) \underset{(1.13)}{=} n(n-1)\,\dots\,(n-k+1),$$

т. е. имеет место формула (1.10). Теорема доказана. Очевидно, что, согласно методу математической индукции, достаточно было проверить справедливость равенства (1.11) и показать, что из формулы (1.13) вытекает формула (1.10).

Определение 6. *Группы, состоящие из одного и того же числа элементов и отличающиеся друг от друга только порядком элементов, называются перестановками.*

Таким образом, перестановки — это размещения из n элементов по n элементов в каждом, $n = 1, 2, \dots$.

Например, группы {1, 2, 3}, {1, 3, 2}, {2, 1, 3}, {2, 3, 1}, {3, 1, 2} и {3, 2, 1} образуют всевозможные перестановки из первых трех натуральных чисел 1, 2, 3.

Число всех перестановок из n элементов обозначается P_n.

ТЕОРЕМА 2. *Имеет место формула* $P_n = 1 \cdot 2 \cdot \dots \cdot n$.

Это сразу следует из формулы (1.10) при $k = n$. Произведение $1 \cdot 2 \cdot \dots \cdot n$ обозначается $n!$ (читается: «эн факториал»). В принятых обозначениях

$$P_n = n! \qquad (1.14)$$

Для удобства полагают $P_0 = 0! = 1$.

Определение 7. *Группы, состоящие из k элементов в каждой и отличающиеся друг от друга по крайней мере одним элементом, называются сочетаниями.*

Например, группы {1, 2}, {1, 3} и {2, 3} образуют все сочетания из натуральных чисел 1, 2, 3 по два из них.

Число всех сочетаний из n элементов по k элементов в каждом обозначается C_n^k.

ТЕОРЕМА 3. *Имеет место формула*

$$C_n^k = \frac{A_n^k}{P_k}, \tag{1.15}$$

т. е.

$$C_n^k = \frac{n(n-1)\dots(n-k+1)}{n!}. \tag{1.16}$$

СЛЕДСТВИЕ. *Справедлива формула*

$$C_n^k = \frac{n!}{k!(n-k)!}. \tag{1.17}$$

Доказательство. Если в каждом сочетании из n элементов по k (их всего C_n^k) сделать всевозможные перестановки его элементов (число таких перестановок равно P_k), то получатся размещения из n элементов по k, причем таким способом получаются все размещения из n элементов по k, и притом по одному разу. Поэтому

$$C_n^k P_k = A_n^k,$$

откуда и следует формула (1.15). Формула (1.16) получается из формул (1.10), (1.14), (1.15).

Если умножить числитель и знаменатель дроби, стоящей в правой части формулы (1.16), на $(n-k)!$, то получится формула (1.17).

ТЕОРЕМА 4. *Имеет место формула*

$$C_n^k = C_n^{n-k},\ k = 0, 1, 2, \dots, n, \tag{1.18}$$

где $C_n^0 \overset{\text{def}}{=} 1$.

Доказательство. Формулу (1.18) легко получить непосредственно из определения сочетаний: если из n элементов выбрать какую-либо группу (сочетание), состоящую из k элементов, то останется группа (сочетание) из $n-k$ элементов, при этом таким способом получаются все сочетания из n элементов по $n-k$ элементов и по одному разу. Поэтому чис-

ло сочетаний из n элементов по k, т. е. C_n^k, равно числу сочетаний из n элементов по $n - k$, т. е. C_n^{n-k}.

Формула (1.18) следует сразу и из равенства (1.17), в силу которого

$$C_n^k = \frac{n!}{k!(n-k)!}, \quad C_n^{n-k} = \frac{n!}{(n-k)!k!},$$

т. е. числа C_n^k и C_n^{n-k} равны.

ТЕОРЕМА 5. *Имеет место формула*

$$C_{n+1}^{k+1} = C_n^k + C_n^{k+1}. \qquad (1.19)$$

Доказательство. Пусть дан $n + 1$ элемент, из которых составляются сочетания по $k + 1$ элементов. Зафиксируем один из этих элементов; тогда число сочетаний, в которые вошел этот элемент, равно C_n^k (так как если его отбросить в каждом таком сочетании, то получатся всевозможные сочетания из n элементов по k элементов и по одному разу), а число сочетаний, в которые он не входит, равно C_n^{k+1} (ибо они образуют всевозможные сочетания по $k + 1$ элементов из оставшихся n элементов). Это и доказывает формулу (1.19).

Замечание. Числа C_n^k можно находить с помощью следующей треугольной таблицы, называемой *треугольником Паскаля*[1], где первые и последние числа во всех строчках равны 1, и начиная с третьей строчки каждое число, отличное от первого и последнего, получается сложением двух ближайших к нему чисел предшествующей строчки:

$$\begin{array}{c}
1 \\
1 \quad 1 \\
1 \quad 2 \quad 1 \\
1 \quad 3 \quad 3 \quad 1 \\
1 \quad 4 \quad 6 \quad 4 \quad 1 \\
\dots\dots\dots\dots\dots\dots \\
1 \dots\, C_n^k \qquad C_n^{k+1} \dots\, 1 \\
1 \dots\dots\dots\, C_{n+1}^{k+1} \dots\dots\dots\, 1 \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots
\end{array}$$

[1] Б. Паскаль (1623—1662) — французский математик.

Из равенств $C_n^0 = C_n^n = 1$ и формулы (1.19) следует, что в n-й строчке этой таблицы стоят числа

$$1,\ C_n^1,\ \dots,\ C_n^k,\ \dots,\ C_n^{n-1},\ 1.$$

1.5. Логические символы

В математических рассуждениях часто встречаются выражения «существует элемент», обладающий некоторыми свойствами, и «любой элемент» среди элементов, имеющих некоторое свойство. Вместо слова «существует» или равносильного ему слова «найдется» иногда пишут символ $\exists$, т. е. перевернутую латинскую букву E (от англ. Existence — существование), а вместо слов «любой», «каждый», «всякий» — символ $\forall$, т. е. перевернутое латинское A (от англ. Any — любой). Символ $\exists$ называется символом *существования*, а символ $\forall$ — символом *всеобщности*.

Примеры. 1. Определение объединения $\bigcup\limits_{\alpha \in \mathfrak{A}} A_\alpha$ множеств A_α, $\alpha \in \mathfrak{A}$, записывается с помощью логического символа существования следующим образом:

$$\bigcup_{\alpha \in \mathfrak{A}} A_\alpha = \{x : \exists\ \alpha \in \mathfrak{A},\ x \in A_\alpha\},$$

а определение пересечения $\bigcup\limits_{\alpha \in \mathfrak{A}} A_\alpha$, записанное с помощью символа всеобщности, имеет вид

$$\bigcup_{\alpha \in \mathfrak{A}} A_\alpha = \{x : \forall\ \alpha \in \mathfrak{A},\ x \in A_\alpha\}.$$

2. Пусть $\boldsymbol{R}$ — множество действительных чисел и пусть задана функция $f : \boldsymbol{R} \to \boldsymbol{R}$, т. е. функция, определенная на множестве действительных чисел и принимающая действительные значения. Функция f называется *четной функцией*, если для любого $x \in \boldsymbol{R}$ выполняется равенство $f(-x) = f(x)$. Используя логическую символику, это условие можно записать короче:

$$\forall\ x \in \boldsymbol{R} : f(-x) = f(x).$$

3. Функция $f : \boldsymbol{R} \to \boldsymbol{R}$ называется *периодической*, если существует такое число $T > 0$, что, каково бы ни было $x \in \boldsymbol{R}$, спра-

ведливо равенство $f(x + T) = f(x)$. Употребляя логические символы, это определение можно записать следующим образом:

$$(\exists\ T > 0)(\forall\ x \in \boldsymbol{R}) : f(x + T) = f(x).$$

Обычно для удобства чтения утверждений, записанных с помощью нескольких логических символов, все, что относится к каждому из них отдельно, заключается в круглые скобки, как это и сделано в последней формуле. Двоеточие в подобных формулах означает «имеет место». Часто подобные выражения записывают для краткости без скобок (мы будем поступать так же):

$$\exists\ T > 0 \quad \forall\ x \in \boldsymbol{R} : f(x + T) = f(x).$$

4. Функция $f : \boldsymbol{R} \to \boldsymbol{R}$ не является четной, если условие $f(-x) = f(x)$ не выполняется для всех $x \in \boldsymbol{R}$. Однако подобные отрицательные формулировки не очень удобны, когда приходится их использовать, так как трудно делать выводы из того, чего нет. Гораздо удобнее иметь дело с позитивными, как их называют, утверждениями, которые не содержат отрицаний. В данном случае утверждение, что равенство $f(-x) = f(x)$ не выполняется для всех $x \in \boldsymbol{R}$, равносильно утверждению, что существует такое $x \in \boldsymbol{R}$, что $f(-x) \neq f(x)$, или в символической записи

$$\exists\ x \in \boldsymbol{R} : f(-x) \neq f(x).$$

5. Функция $f : \boldsymbol{R} \to \boldsymbol{R}$ не является периодической, если любое число $T > 0$ не является ее периодом, т. е. для любого $T > 0$ равенство $f(x + T) = f(x)$ не должно выполняться для всех $x \in \boldsymbol{R}$, или в позитивной форме: для любого $T > 0$ найдется $x \in \boldsymbol{R}$, для которого $f(x + T) \neq f(x)$. С помощью логических символов это определение записывается следующим образом:

$$\forall\ T > 0 \quad \exists\ x \in \boldsymbol{R} : f(x + T) \neq f(x).$$

Сравнив запись с помощью логических символов утверждений в примерах 2 и 3 с их отрицанием в примерах 4 и 5, видим, что при построении отрицаний символы существования и всеобщности заменяют друг друга. Для того чтобы в некотором множестве не существовал элемент, обладающий каким-то свойством, надо, чтобы все элементы не обладали этим свойством, т. е. в этом случае при отрицании символ существования $\exists$ переходит в символ всеобщности $\forall$. Если

же каким-то свойством обладают не все элементы рассматриваемого множества, то это означает, что в нем существует элемент, не обладающий данным свойством: символ всеобщности заменился символом существования.

Символ $\Rightarrow$ означает «следует» (одно высказывание следует из другого), а символ $\Leftrightarrow$ означает равносильность высказываний, стоящих по разные от него стороны. Значок def означает, что сформулированное утверждение справедливо по определению (от англ. definition — определение). Например,

$$A \subset B \overset{\text{def}}{\Leftrightarrow} (\forall\, x \in A \Rightarrow x \in B), \quad (g \circ f)(x) \overset{\text{def}}{=} g(f(x)).$$

Определение часто используемого в математике символа Σ (греческая заглавная буква «сигма») для обозначения суммы слагаемых можно записать следующим образом:

$$\sum_{k=1}^{n} a_k \overset{\text{def}}{=} a_1 + a_2 + \ldots + a_n.$$

Как правило, изложение материала ведется в классическом стиле без использования логических символов. Они употребляются лишь параллельно с основным текстом. Это, с одной стороны, поможет читателю привыкнуть к их применению, что, например, полезно при конспектировании, а с другой — более кратко и, следовательно, иногда более выразительно разъяснить нужную мысль.

В дальнейшем конец проводимого доказательства сформулированного утверждения будет отмечаться символом $\square$.

§ 2.

Действительные числа

2.1. Свойства действительных чисел

В элементарной математике изучаются действительные (вещественные) числа. Сначала в процессе счета возникает так называемый *натуральный ряд* чисел 1, 2, 3, ... , n, ... В арифметике вводятся действия сложения и умножения над натуральными числами. Что же касается операций вычитания и деления, то они уже оказываются не всегда возможными в множестве натуральных чисел. Чтобы все четы-

ре арифметические операции были возможны для любой пары чисел (кроме операции деления на нуль, которой нельзя приписать смысла), приходится расширить класс рассматриваемых чисел. К необходимости такого расширения запаса чисел приводят также потребности измерения тех или иных геометрических и физических величин. Поэтому вводятся число нуль и целые *отрицательные* числа (вида $-1, -2, \ldots, -n, \ldots$), а затем и *рациональные* (вида p/q, где p, q — целые, $q \neq 0$).

Та же потребность измерения величин и проведения таких операций, как извлечение корня, вычисление логарифмов, решение алгебраических уравнений, приводит к дальнейшему расширению запаса рассматриваемых чисел: появляются *иррациональные* и, наконец, *комплексные числа*. Все рациональные и все иррациональные числа образуют множество всех действительных чисел.

Множество всех действительных чисел, как принято, будем обозначать через $\boldsymbol{R}$ (от лат. realus — действительный). Это множество образует совокупность, в которой определены взаимосвязные операции сложения, умножения и сравнения чисел по величине и которая обладает определенного рода непрерывностью. Напомним кратко свойства действительных чисел, известные из элементарной математики, и дополним их описанием некоторых свойств, обычно не рассматриваемых там достаточно полно.

I. Операция сложения. Для любой упорядоченной пары действительных чисел a и b определено, и притом единственным образом, число, называемое их *суммой* и обозначаемое через $a + b$, так что при этом имеют место следующие свойства.

I_1. *Для любой пары чисел a и b*

$$a + b = b + a.$$

Это свойство называется *переместительным* или *коммутативным законом сложения*.

I_2. *Для любых чисел a, b и c*

$$a + (b + c) = (a + b) + c.$$

Это свойство называется *сочетательным* или *ассоциативным законом сложения*.

I_3. *Существует число, обозначаемое* 0 *и называемое нулем, такое, что для любого числа* a

$$a + 0 = a.$$

I_4. *Для любого числа* a *существует число, обозначаемое* $-a$ *и называемое противоположным данному, такое, что*

$$a + (-a) = 0.$$

II. Операция умножения. Для любой упорядоченной пары чисел a и b определено, и притом единственным образом, число, называемое их *произведением* и обозначаемое ab (или $a \cdot b$), так что при этом имеют место следующие свойства.

II_1. *Для любой пары чисел* a *и* b

$$ab = ba.$$

Это свойство называется *переместительным* или *коммутативным законом умножения.*

II_2. *Для любых чисел* a, b, c

$$a(bc) = (ab)c.$$

Это свойство называется *сочетательным* или *ассоциативным законом умножения.*

II_3. *Существует число, обозначаемое* 1 *и называемое единицей, такое, что для любого числа* a

$$a \cdot 1 = a.$$

II_4. *Для любого числа* $a \neq 0$ *существует число, обозначаемое* $1/a$ *или* $\frac{1}{a}$ *и называемое обратным данному, такое, что*

$$a \cdot \frac{1}{a} = 1.$$

III. Связь операций сложения и умножения. *Для любых чисел* a, b, c

$$(a + b)c = ac + bc.$$

Это свойство называется *распределительным* или *дистрибутивным законом умножения* относительно сложения.

IV. Упорядоченность. *Для каждого числа* a *определено одно из соотношений* $a > 0$ (*a больше нуля*), $a = 0$

(*а равно нулю*) *или* $a < 0$ (*а меньше нуля*) *так, что условие* $a > 0$ *равносильно условию* $-a < 0$.

При этом если $a > 0$, $b > 0$, *то имеют место неравенства*:

Рис. 2

IV_1. $a + b > 0$.

IV_2. $ab > 0$.

Свойство IV дает возможность ввести понятие сравнения или, как иногда говорят, сравнения по величине для любых двух чисел.

Число b называют числом, бо́льшим числа a, и пишут $b > a$, или, что то же самое, число a называют меньшим числа b и пишут $a < b$, если $b - a > 0$.

Наличие сравнения «больше» или «меньше» для любой пары действительных чисел называется *свойством упорядоченности множества всех действительных чисел*.

Действительные числа обладают еще так называемым свойством непрерывности.

V. Свойство непрерывности. *Каковы бы ни были непустые множества* $A \subset \boldsymbol{R}$ *и* $B \subset \boldsymbol{R}$, *у которых для любых двух элементов* $a \in A$ *и* $b \in B$ *выполняется неравенство* $a \leqslant b$, *существует такое число* α, *что для всех* $a \in A$ *и* $b \in B$ *имеет место* (рис. 2) *соотношение*

$$a \leqslant \alpha \leqslant b.$$

Свойство непрерывности действительных чисел связано с самым простейшим использованием математики на практике — с измерением величин. При измерении какой-либо физической или какой-нибудь другой природы величины часто получают с большей или меньшей точностью ее приближенные значения. Если в результате экспериментального измерения данной величины получается ряд чисел, дающих значение искомой величины с недостатком (они играют роль множества A в приведенной выше формулировке свойства непрерывности) и с избытком (множество B), то свойство непрерывности действительных чисел выражает объективную уверенность в том, что измеряемая величина имеет определенное значение, расположенное между ее приближенными значениями, вычисленными с недостатком и избытком.

Из свойств I—V действительных чисел вытекают другие многочисленные их свойства, поэтому можно сказать, что действительные числа представляют собой совокупность элементов, обладающую свойствами I—V.

Для вдумчивого читателя заметим, что ссылка в начале параграфа на то, что действительные числа и их свойства известны из курса элементарной математики, не является необходимой. Сформулированные выше свойства действительных чисел можно взять за исходное определение. Следует только исключить тривиальный случай: легко проверить, что для множества, состоящего только из одного нуля, выполняются все свойства I—V (в таком множестве $0 = 1$). Множество, в котором имеется хоть один элемент, отличный от нуля, будем здесь для краткости называть нетривиальным.

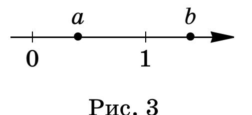


Рис. 3

Перефразируя сказанное, получим следующее определение.

Определение 1. *Нетривиальное множество элементов, обладающих свойствами* I—V, *называется множеством действительных чисел. Каждый элемент этого множества называется действительным числом.*

Построение теории действительных чисел, основывающееся на таком их определении, называется *аксиоматическим*, а свойства I—V — *аксиомами действительных чисел*.

Геометрически множество действительных чисел изображается направленной (ориентированной) прямой, а отдельные числа — точками этой прямой. Поэтому совокупность действительных чисел часто называют числовой прямой, а также числовой или действительной осью, а отдельные числа — ее точками (рис. 3).

При такой интерпретации действительных чисел иногда вместо a меньше b (a больше b) говорят, что точка a лежит левее точки b (что a лежит правее b).

В п. 2.2*—2.6* будут более детально проанализированы свойства I—V действительных чисел и выведены некоторые их следствия. Как и все пункты, отмеченные звездочками, эти пункты при первом чтении можно без существенного ущерба для усвоения курса математического анализа опустить. Для понимания дальнейшего материала (в § 3 и следующих) вполне достаточно представления о действительных числах, которое дается в курсе элементарной математики.

2.2*. Свойства сложения и умножения

Рассмотрим некоторые свойства сложения и умножения, которые вытекают из свойств I, II и III. Прежде всего заметим,

что для операции сложения существует обратная операция — вычитание; определим ее.

Для любой упорядоченной пары чисел $a \in \boldsymbol{R}$ и $b \in \boldsymbol{R}$ число $a + (-b)$ называется *разностью* чисел a и b и обозначается через $a - b$, т. е.

$$a - b \stackrel{\text{def}}{=} a + (-b). \tag{2.1}$$

Если

$$a + b = c, \tag{2.2}$$

то, прибавляя к обеим частям этого равенства число $-b$, получаем $(a + b) + (-b) = c + (-b)$. Отсюда, согласно ассоциативному закону I_2 и определению разности, имеем

$$a + (b + (-b)) = c - b,$$

но $b + (-b) = 0$; следовательно,

$$a = c - b. \tag{2.3}$$

Таким образом, после прибавления к числу a числа b число a восстанавливается вычитанием из суммы $a + b$ числа b, поэтому операция вычитания и называется *операцией, обратной операции сложения*.

Перейдем теперь к свойствам сложения и умножения действительных чисел.

1^0. *Число, обладающее свойством нуля, единственно.*

Действительно, допустим, что существуют два нуля 0 и $0'$; тогда, в силу I_3: $0' + 0 = 0'$, $0 + 0' = 0$. Согласно коммутативному закону I_2, левые части этих равенств равны, следовательно, равны и правые, т. е. $0 = 0'$. □

2^0. *Число, противоположное данному, единственно.*

Пусть числа b и c противоположны некоторому числу a, т. е. $a + b = 0$ и $a + c = 0$. Тогда из первого равенства имеем $(a + b) + c = 0 + c$, т. е. $(a + b) + c = c$, откуда $(a + c) + b = c$; но $a + c = 0$, следовательно, $b = c$. □

3^0. *Для любого числа a справедливо равенство*

$$-(-a) = a.$$

Из равенства $a + (-a) = 0$, определяющего противоположный элемент, в силу коммутативности сложения, получим $(-a) + a = 0$. Это и означает, что $a = -(-a)$. □

4^0. *Для любого числа a справедливо равенство*

$$a - a = 0.$$

В самом деле, $a - a = a + (-a) = 0$. □

5^0. *Для любых чисел a и b имеем*

$$-a - b = -(a + b),$$

т. е. число, противоположное сумме двух чисел, равно сумме противоположных им чисел.

Действительно, $a + b + (-a - b) = (a - a) + (b - b) = 0$. □

6^0. *Уравнение $a + x = b$ имеет в $\boldsymbol{R}$ решение, и притом единственное*: $x = b - a$.

В самом деле, если решение существует, то, в силу формул (2.2)—(2.3), $x = b - a$. Этим и доказана единственность решения уравнения $a + x = b$. Для существования решения достаточно проверить, что число $x = b - a$ является решением. Это действительно так:

$$a + (b - a) = a + [b + (-a)] = [a + (-a)] + b = b. \ \square$$

Для операции умножения также существует обратная операция; она называется делением и определяется следующим образом.

Для любой упорядоченной пары чисел a и b, $b \neq 0$, число $a \cdot \frac{1}{b}$ называется *частным от деления* a на b и обозначается через $\frac{a}{b}$, или a/b, или $a : b$, т. е.

$$\frac{a}{b} \overset{\text{def}}{=} a \cdot \frac{1}{b}, \quad b \neq 0.$$

Запись частного от деления a на b в виде $\frac{a}{b}$ (или a/b) называется дробью с числителем a и знаменателем b.

Свойства, аналогичные свойствам 1^0—6^0 для сложения, справедливы и для операции умножения:

7^0. *Число, обладающее свойствами единицы, единственно.*

8^0. *Число, обратное данному числу, отличному от нуля, единственно.*

9^0. *Для любого числа $a \neq 0$ справедливо равенство*

$$\frac{1}{1/a} = a.$$

10^0. *Для любого числа $a \neq 0$ справедливо равенство*

$$\frac{a}{a} = 1.$$

11^0. *Для любых чисел $a \neq 0$ и $b \neq 0$ имеем равенство*

$$\frac{1}{a} \cdot \frac{1}{b} = \frac{1}{ab},$$

т. е. число, обратное произведению чисел, отличных от нуля, равно произведению обратных им чисел.

12^0. *Уравнение $ax = b$, $a \neq 0$, имеет в множестве действительных чисел и притом единственное решение $x = \frac{b}{a}$.*

Свойства 7^0—12^0 доказываются аналогично свойствам 1^0—6^0.

13^0. *Равенство* $\frac{a}{b} = \frac{c}{d}$, $b \neq 0$, $d \neq 0$, справедливо тогда и только тогда, когда $ad = bc$.

СЛЕДСТВИЕ **(основное свойство дроби).** *Каковы бы ни были дробь a/b, $b \neq 0$, и число $c \neq 0$, имеет место равенство*

$$\frac{a}{b} = \frac{ac}{bc}.$$

Действительно, умножая обе части равенства $a/b = c/d$ на bd и используя определение деления, имеем следующую цепочку эквивалентных равенств:

$$\frac{a}{b} = \frac{c}{d} \Leftrightarrow \frac{a}{b}bd = \frac{c}{d}db \Leftrightarrow a \cdot \frac{1}{b}bd = c\frac{1}{d}db \Leftrightarrow ad = cb. \ \square$$

Все рассмотренные свойства 1^0—13^0 касаются только операций сложения и умножения. Эти операции позволяют определить натуральные, целые и рациональные числа, операцию возведения в целую степень и операцию извлечения корня. Проделаем это.

Число $1 + 1$ обозначается через 2, число $2 + 1$ — через 3 и т. д. Числа 1, 2, 3, ... называются *натуральными числами*. Их обозначение и название совпадают с числами элементов в конечных множествах (см. п. 1.3*). Это не случайно, поскольку, для того чтобы получить натуральное число n в новом смысле, надо взять конечное множество единиц, число элементов которого в п. 1.3* было обозначено тем же символом n, и сложить их (см. (1.4)). При этом отношение поряд-

ка, введенное в множестве натуральных чисел (см. п. 1.3*), совпадает с порядком, имеющимся в этом множестве согласно упорядоченности множества всех действительных чисел (см. свойство IV в п. 2.1), причем натуральным числом, непосредственно следующим за n, является $n + 1$. Как уже отмечалось, множество натуральных чисел обозначается через $\boldsymbol{N}$.

Заметим, что, хотя, как это было доказано выше, единица единственна, можно рассматривать несколько экземпляров единицы (как и вообще несколько экземпляров любого элемента некоторого множества) хотя бы для того, чтобы можно было написать выражение $1 + 1$.

Числа $0, \pm 1, \pm 2, \ldots$ называются *целыми числами*. Множество целых чисел обычно обозначается через $\boldsymbol{Z}$.

В дальнейшем будет показано (см. свойство 8^0 в п. 2.3*), что из всех перечисленных в п. 2.1 свойств действительных чисел следует: $1 > 0$.

Числа вида m/n, где m и n — целые, а $n \neq 0$, называются *рациональными числами*. Множество всех рациональных чисел обозначается обычно через $\boldsymbol{Q}$. Действительные числа, не являющиеся рациональными, называются *иррациональными*. Их множество обозначается через $\boldsymbol{I}$.

Отметим теперь несколько свойств, которые связывают операции сложения и умножения.

14^0. *Для любых чисел a, b и c имеет место равенство*

$$a(b - c) = ab - ac.$$

В самом деле,

$$a(b - c) = a(b - c) + ac - ac = a(b - c + c) - ac = ab - ac. \ \square$$

15^0. *Для любого числа a выполняется равенство*

$$a \cdot 0 = 0.$$

Действительно, возьмем какое-либо число b; тогда $b - b = 0$ (см. свойство 4^0). Поэтому, согласно свойству 14^0, имеем

$$a \cdot 0 = a(b - b) = ab - ab = 0. \ \square$$

Из свойства 14^0, между прочим, вытекает, что утверждение $1 \neq 0$ при выполнении свойств I—III эквивалентно тому, что существует хотя бы одно число, отличное от нуля. Очевидно, достаточно показать, что если существует число $a \neq 0$, то $1 \neq 0$. Докажем это: пусть существует $a \neq 0$; тогда из равен-

ства $a \cdot 1 = a$ следует, что $1 \neq 0$, так как в противном случае, согласно свойству 15^0, имело бы место равенство $a = 0$.

16^0. *Если $ab = 0$, то по крайней мере один из сомножителей a и b равен нулю.*

Пусть, например, $a \neq 0$; тогда, умножив равенство $ab = 0$ на $\frac{1}{a}$, получим $\frac{1}{a}(ab) = \frac{1}{a} \cdot 0$, откуда $\left(\frac{1}{a}a\right)b = 0$, следовательно, $b = 0$. □

17^0. *Для любых чисел a и b имеем*

$$(-a)b = -ab,\ (-a)(-b) = ab;$$

в частности, $(-1)a = -a$.

В самом деле,

$$(-a)b = (-a)b + ab - ab = (-a + a)b - ab = -ab.$$

Отсюда, в силу коммутативности умножения (меняя a и b местами), имеем $a(-b) = -ab$. Используя это равенство, получаем

$$(-a)(-b) = -a(-b) = (-1)[a(-b)] = (-1)(-ab) = -(-ab) = ab. \ \square$$

Из свойств I, II, III действительных чисел и перечисленных выше их следствий можно получить правила арифметических действий с дробями, т. е. числами вида a/b, $b \neq 0$, $a \in \boldsymbol{R}$, $b \in \boldsymbol{R}$.

18^0. *Сложение дробей производится по правилу*

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd},\ b \neq 0,\ d \neq 0.$$

Докажем справедливость этого равенства. Использовав определение деления, дистрибутивность сложения относительно умножения и основное свойство дроби, получим

$$\frac{ad + bc}{bd} = (ad + bc)\frac{1}{bd} = ad\frac{1}{bd} + bc\frac{1}{bd} = \frac{ad}{bd} + \frac{bc}{bd} = \frac{a}{b} + \frac{c}{d}. \ \square$$

19^0. *Умножение дробей производится по правилу*

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd},\ b \neq 0,\ d \neq 0.$$

Использовав определение деления и свойство 11^0, получим

$$\frac{ac}{bd} = ac\frac{1}{bd} \underset{(11^0)}{=} ac\frac{1}{b}\frac{1}{d} = \left(a \cdot \frac{1}{b}\right)\left(c\frac{1}{d}\right) = \frac{a}{b}\frac{c}{d}.$$

20^0. *Обратным элементом дроби $\frac{a}{b}$, $a \neq 0$, $b \neq 0$ является дробь $\frac{b}{a}$, т. е. $\frac{a}{b} \cdot \frac{b}{a} = 1$.*

Это сразу следует из правила умножения дробей.

21^0. *Деление дробей производится по правилу*

$$\frac{a}{b} : \frac{c}{d} = \frac{ad}{bc},\ b \neq 0,\ c \neq 0,\ d \neq 0.$$

Использовав определение деления, предыдущее свойство и правило умножения дробей, будем иметь

$$\frac{a}{b} : \frac{c}{d} = \frac{a}{b} \cdot \frac{1}{\frac{c}{d}} = \frac{a}{b} \cdot \frac{d}{c} = \frac{ad}{bc}. \ \square$$

Пусть заданы действительное число a и натуральное число n. Число a, умноженное n раз на себя, называется n-й *степенью* числа a и обозначается через a^n. Таким образом,

$$a^n \overset{\text{def}}{=} \underbrace{aa \ldots a}_{n \text{ раз}}.$$

По определению полагается $a^0 \overset{\text{def}}{=} 1$ и для любого $n \in \boldsymbol{N}$ $a^{-n} \overset{\text{def}}{=} \frac{1}{a^n}$.

Выведем теперь из полученных выше свойств сложения и умножения действительных чисел правила действий со степенями.

22^0. *Если m и n — целые числа, причем в случае, когда $m \leqslant 0$ или $n \leqslant 0$, имеет место $a \neq 0$, то*

$$a^m a^n = a^{m+n},\ (a^m)^n = a^{mn}.$$

Если $m = 0$ или $n = 0$, то справедливость формул очевидна. В том случае, когда m и n — натуральные числа, согласно определению степени,

$$a^m a^n = \underbrace{a \ldots a}_{m \text{ раз}}\ \underbrace{a \ldots a}_{n \text{ раз}} = a^{m+n}.$$

Если $m < 0$, $n > 0$ и $a \neq 0$, то, полагая $k = -m$ и используя основное свойство дроби (возможность одновременного деления числителя и знаменателя дроби на одно и то же не равное нулю число без нарушения равенства), при $k \leqslant n$ будем иметь

$$a^m a^n = a^{-k} a^n = \frac{a^n}{a^k} = \frac{\overbrace{a \ldots a}^{n \text{ раз}}}{\underbrace{a \ldots a}_{k \text{ раз}}} = a^{n-k} = a^{m+n},$$

а при $k > n$

$$a^m a^n = \frac{a^n}{a^k} = \frac{1}{a^{k-n}} = a^{n-k} = a^{m+n}.$$

Если $m < 0$, $n < 0$ и $a \neq 0$, то, полагая, $k = -m$, $l = -n$ и используя свойство 11^0, получаем

$$a^m a^n = a^{-k} a^{-l} = \frac{1}{a^k}\frac{1}{a^l} = \frac{1}{a^{k+l}} = a^{-(k+l)} = a^{m+n}.$$

Аналогично проверяется и вторая формула свойства 22^0. □

Легко показать, что свойства I_1, I_2, II_1, II_2 и III распространяются по индукции на любое конечное число членов. В качестве примера покажем, что для любых чисел $a_1, a_2, \ldots, a_n$ $(n \geqslant 2)$ и b

$$(a_1 + a_2 + \ldots + a_n)b = a_1 b + a_2 b + \ldots + a_n b. \qquad (2.4)$$

В самом деле, при $n = 2$ эта формула справедлива согласно свойству III.

Пусть теперь формула (2.4) справедлива при $n = k$, покажем, что она будет справедлива и при $n = k + 1$. Применив свойство I_2 для $k + 1$ слагаемых (считая, что оно уже доказано), затем свойство III и использовав предположение индукции, получим

$$(a_1 + a_2 + \ldots + a_{k+1})b = [(a_1 + \ldots + a_k) + a_{k+1}]b =$$
$$= (a_1 + \ldots + a_k)b + a_{k+1}b = a_1 b + \ldots + a_k b + a_{k+1} b.$$

Из формулы (2.4) в случае $a_1 = a_2 = \ldots = a_n = 1$ следует, что

$$nb = \underbrace{b + \ldots + b}_{n \text{ раз}},$$

т. е. что умножение числа на натуральное число n сводится к сложению этого числа n раз.

З а м е ч а н и е. Отметим, что свойства I—III п. 2.1 не описывают полностью действительные числа в том смысле, что существуют и другие множества, отличные от совокупности действительных чисел, удовлетворяющие тем же свойствам I—III, если в них слово «число» всюду заменить словом «элемент» рассматриваемого множества. Именно в этом смысле всюду в дальнейшем понимается выражение «множество, удовлетворяющее каким-либо из свойств I—III».

Примером множеств, удовлетворяющих условиям I, II и III, являются одни только рациональные числа или известные из элементарной математики комплексные числа, а также совокупность рациональных функций, т. е. функций вида $f(x) = \frac{P(x)}{Q(x)}$, где $P(x)$ и $Q(x)$ — многочлены.

Элементы всех перечисленных множеств можно складывать и умножать, причем эти операции будут подчиняться условиям I, II и III. Множества, удовлетворяющие этим требованиям и содержащие хотя бы один элемент, отличный от нуля, называются *полями*.

Таким образом, рациональные числа, действительные числа, комплексные числа и рациональные функции образуют поля.

Проанализируем теперь свойства, выделяющие поле действительных чисел среди всех других полей. Одним из таких свойств является свойство упорядоченности его элементов.

2.3*. Свойства упорядоченности

Выведем некоторые следствия из свойств упорядоченности IV и свойств сложения и умножения I, II и III.

Прежде всего напомним понятие сравнения по величине для любых двух чисел при условии, что для любого числа a определено, меньше ли оно нуля, равно ли нулю или больше его, причем никакое число и противоположное ему не могут быть одновременно больше или, соответственно, меньше нуля (в этом состояло свойство IV): число b называется числом, большим числа a: $b > a$, если $b - a > 0$.

Соотношение $a < b$ или равносильное ему $b > a$ называются *неравенствами* (иногда строгими неравенствами).

Заметим, что соотношение $a < b$ в случае $a = 0$ или $b = 0$ совпадает с исходным соотношением IV. Действительно, если, например, $b = 0$ и $a < 0$ в первоначальном смысле, то, согласно свойству IV, имеет место неравенство $-a > 0$. Но $-a = 0 - a$, поэтому $0 - a > 0$ и, следовательно, $a < 0$ и в смысле последующего определения. Обратно: если $a < 0$ в смысле последующего определения, т. е. $0 - a > 0$, следовательно, $-a > 0$, то, согласно свойству IV, это равносильно условию $a < 0$ в первоначальном смысле.

Рассмотрим теперь основные свойства сравнения чисел по величине.

1^0. *Если* $a > b$ *и* $b > c$, *то* $a > c$.

Это свойство называется т р а н з и т и в н о с т ь ю упорядоченности (сравнения по величине) чисел.

Если $a > b$ и $b > c$, то, согласно определению, это означает, что $a - b > 0$ и $b - c > 0$. Складывая эти неравенства, согласно IV_1 получаем $(a - b) + (b - c) > 0$, т. е. $a - c > 0$. Это и означает, что $a > c$. □

2^0. *Если* $a > b$, *то для любого числа* c *имеем* $a + c > b + c$.

В самом деле, неравенство $a > b$ означает, что $a - b > 0$. Так как, согласно свойству 5^0 из п. 2.2^*, $a - b = a + c - c - b = (a + c) - (b + c)$, то $(a + c) - (b + c) > 0$ и, следовательно, $a + c > b + c$. □

Соотношение $a < b$ читается: «a меньше b». Соотношение $a = b$ читается: «a равно b». Соотношение $a > b$ читается: «a больше b».

Наличие транзитивного отношения порядка «больше», «меньше» между любыми двумя числами называется обычно *свойством упорядоченности множества действительных чисел* или *отношением порядка*.

Запись $a \leqslant b$ равнозначна записи $b \geqslant a$ и означает, что либо $a = b$, либо $a < b$. Например, можно написать $2 \leqslant 2$, $2 \leqslant 5$. Конечно, можно написать более точно: $2 = 2$, $2 < 5$, однако неравенства $2 \leqslant 2$ и $2 \leqslant 5$ также верны, так как означают, что «два не больше двух», соответственно что «два не больше пяти».

3^0. *Для любых двух чисел* a *и* b *имеется в точности одно из трех отношений порядка* $a > b$, $a = b$ *или* $a < b$.

Действительно, пусть заданы два числа a и b. Для их разности $a - b$, согласно свойству IV, имеет место точно одно из соотношений $a - b > 0$, $a - b = 0$ или $0 > a - b$.

Если $a - b > 0$, то, по определению, $a > b$. Если $a - b = 0$, то, прибавив к обеим частям равенства число b, получим $a = b$. Наконец, если $0 > a - b$, то, прибавив последовательно к обеим частям неравенства $0 > a - b$ числа $-a$ и b (см. предыдущее свойство), получим $b - a > 0$. Это и означает, что $b > a$, или, что то же, $a < b$. □

4^0. *Если* $a < b$, *то* $-a > -b$.

Действительно, из $a < b$ в силу определения имеем $b - a > 0$. Поэтому $-a = -a + b + (-b) = (b - a) + (-b) > 0 + (-b) = -b$. □

5^0. *Если* $a < b$ *и* $c \leqslant d$, *то* $a + c < b + d$, *т. е. можно производить почленное сложение неравенств одного знака.*

В самом деле, если $a < b$ и $c \leqslant d$, то, согласно свойству 2^0 этого пункта, $a + c < b + c$ и $c + b \leqslant d + b$, поэтому, в силу транзитивности упорядоченности, имеем $a + c < b + d$. □

6^0. *Если* $a < b$ *и* $c \geqslant d$, *то* $a - c < b - d$, *т. е. неравенства противоположных знаков можно вычитать в указанном смысле.*

Действительно, из $c \geqslant d$ имеем, согласно свойству 4^0 этого пункта, $-c \leqslant -d$. Сложив неравенства $a < b$ и $-c \leqslant -d$, получим $a - c < b - d$. □

7^0. *Если* $a < b$ *и* $c < 0$, *то* $ac > bc$.

В самом деле, согласно свойству 4^0 этого пункта, $-c > 0$, поэтому, в силу свойства IV_2, имеем $a(-c) < b(-c)$. Отсюда, согласно свойству 17^0 п. 2.2^*, получим $-ac < -bc$ и, следовательно (см. свойства 4^0 этого пункта), $ac > bc$. □

Из свойства 7^0 следует, что если $b > 0$, $d > 0$, то условие $\frac{a}{b} < \frac{c}{d}$ равносильно условию $ad < bc$. В самом деле, второе из неравенств получается из первого умножением обеих его частей на bd, а первое из второго — делением на bd.

Из свойства 7^0 (при $a \neq 0$) и из свойства IV_2 вытекает правило знаков при умножении действительных чисел: *произведение двух сомножителей одного знака* (либо одновременно положительных, либо одновременно отрицательных) *положительно, а произведение двух сомножителей разных знаков* (один из них положителен, другой — отрицателен) *отрицательно.*

8^0. *В упорядоченном поле всегда справедливо неравенство* $1 > 0$.

В самом деле, как уже было показано (см. замечание после свойства 14^0 в п. 2.2^*), из условия существования элемента $a \neq 0$ (это условие входит в определение поля, см. конец п. 2.2^*) следует, что $1 \neq 0$. Покажем, что неравенство $1 < 0$ невозможно. Допустим противное, пусть $1 < 0$. Возьмем ка-

кое-либо $a > 0$. Согласно определению единицы, имеем $a \cdot 1 = a$. По правилу знаков произведение положительного числа a и отрицательного, по предположению 1, является отрицательным числом, т. е. $a < 0$, что противоречит условию. □

Действительные числа снова не являются единственным объектом, который удовлетворяет аксиомам I—IV. Множества, для которых справедливы эти аксиомы, называются *упорядоченными полями*. Примером упорядоченного поля, отличного от поля действительных чисел, является поле рациональных чисел. Однако уже ни поле комплексных чисел, ни поле рациональных функций не являются упорядоченным полем.

Во всяком упорядоченном поле можно ввести понятие абсолютной величины его элементов. При ее определении и изучении ее свойств для единообразия изложения будем все время говорить о числах, а не об элементах произвольного упорядоченного поля.

Для любого числа a число, обозначаемое $|a|$ и определяемое по формуле

$$|a| = \begin{cases} a, \text{ если } a \geqslant 0, \\ -a, \text{ если } a < 0, \end{cases}$$

называется *абсолютной величиной* числа a, или, что то же, его *модулем*.

Отметим ряд свойств абсолютной величины.

1^0. *Для любого числа a выполняются соотношения*

$$|a| \geqslant 0, \tag{2.5}$$

$$|a| = |-a|, \tag{2.6}$$

$$a \leqslant |a|,\ -a \leqslant |a|. \tag{2.7}$$

Докажем неравенство (2.5). Если $a \geqslant 0$, то $|a| = a \geqslant 0$; если же $a < 0$, то $|a| = -a > 0$.

Докажем равенство (2.6). Если $a \geqslant 0$, то $|a| = a$ и $-a \leqslant 0$, поэтому, согласно определению абсолютной величины и свойству 3 из п. 2.2*, получим $|-a| = -(-a) = a = |a|$. Если же $a < 0$, то $|a| = -a$ и $-a > 0$; это означает, что $|-a| = -a$.

Докажем неравенство (2.7). Если $a \geqslant 0$, то $a = |a|$ и $-a \leqslant \leqslant 0 \leqslant a = |a|$, т. е. (2.7) выполняется. Если же $a < 0$, то $a < 0 < < -a = |a|$, т. е. (2.7) также выполняется.

2^0. *Для любых чисел a и b*

$$|a + b| \leqslant |a| + |b|, \tag{2.8}$$

$$||a| - |b|| \leqslant |a - b|. \tag{2.9}$$

Докажем эти неравенства. Согласно (2.7), имеем:

$$a \leqslant |a|, -a \leqslant |a|, b \leqslant |b|, -b \leqslant |b|.$$

Отсюда, в силу свойства 5^0 из п. 2.3* и свойства 5^0 из п. 2.2*,

$$a + b \leqslant |a| + |b|, -(a + b) \leqslant |a| + |b|.$$

Одно из чисел $a + b$ или $-(a + b)$ неотрицательно и, следовательно, совпадает с $|a + b|$. Неравенство (2.8) доказано.

Неравенство же (2.9) является следствием (2.8). В самом деле,

$$|a| - |b| = |(a - b) + b| - |b| \leqslant |a - b| + |b| - |b| = |a - b|;$$

аналогично,

$$|b| - |a| \leqslant |b - a| = |a - b|.$$

Согласно свойству 5^0 п. 2.2*, $|b| - |a| = -(|a| - |b|)$. Одно из чисел $|a| - |b|$ и $-(|a| - |b|)$ совпадает с $||a| - |b||$. Неравенство (2.9) также доказано. □

3^0. *Для любых чисел a и b выполняется равенство*

$$|ab| = |a||b|.$$

Это сразу следует из определения абсолютной величины, свойства 17^0 п. 2.2* и правила знаков при умножении.

Рассмотрим теперь свойство непрерывности, которое выделяет поле действительных чисел среди всех прочих упорядоченных полей.

2.4*. Свойство непрерывности действительных чисел

Упорядоченное поле, удовлетворяющее свойству V, называется *непрерывным упорядоченным полем*. Поле рациональных чисел уже не является непрерывным упорядоченным полем: в нем имеются множества A и B, для любых элементов $a \in A$ и $b \in B$ которых выполняется неравенство $a < b$, вместе с тем не существует такого рационального числа r,

чтобы для всех $b \in B$ и $a \in A$ выполнялось соотношение $a \leqslant \leqslant r \leqslant b$. Этим свойством обладают, например, множество B, которое состоит из всех положительных рациональных чисел r, удовлетворяющих неравенству $r^2 > 2$, и множество A, в которое отнесены все остальные рациональные числа, так как не существует рационального числа, квадрат которого равнялся бы числу 2.

В терминах упорядоченных полей определение множества действительных чисел можно сформулировать следующим образом.

Определение 2. *Множеством действительных чисел называется непрерывное упорядоченное поле.*

Поле рациональных чисел, как уже отмечалось выше, не обладает свойством непрерывности, а поле действительных чисел — обладает. Поэтому заведомо существуют действительные числа, не являющиеся рациональными, т. е. существуют иррациональные числа. Таким образом, множество действительных чисел является расширением множества рациональных чисел в том смысле, что множество рациональных чисел является собственным подмножеством множества действительных чисел. При этом расширении сохраняются свойство упорядоченности и операции сложения и умножения. Оказывается, что действительные числа, в отличие от рациональных, уже нельзя расширить до большего множества так, чтобы сохранялись указанные свойства (упорядоченность и операции сложения и умножения). Это свойство называется свойством *полноты действительных чисел* относительно их упорядоченности, сложения и умножения. Его доказательство будет дано в п. 3.8*.

2.5*. Сечения в множестве действительных чисел

Свойство непрерывности действительных чисел можно формулировать в различных терминах. Рассмотрим формулировку этого свойства в терминах так называемых с е ч е н и й действительных чисел. Прежде всего определим это понятие.

Определение 3. *Два множества $A \subset \boldsymbol{R}$ и $B \subset \boldsymbol{R}$ называются сечением множества действительных чисел $\boldsymbol{R}$, если*:

1^0) *объединение множеств A и B составляет все множество действительных чисел $\boldsymbol{R}$, $A \cup B = \boldsymbol{R}$;*

2^0) *каждое из множеств A и B не пусто*, $A \neq \varnothing$, $B \neq \varnothing$;

3^0) *каждое число множества A меньше любого числа множества B: если* $a \in A$, $b \in B$, *то* $a < b$.

Свойство 1^0) означает, что каждое действительное число принадлежит по крайней мере одному из множеств A и B.

Из свойства 3^0) очевидно следует, что множества A и B не пересекаются: $A \cap B = \varnothing$. Действительно, если бы нашелся элемент $x \in A \cap B$, т. е. $x \in A$ и $x \in B$, то из свойства 3^0) следовало бы, что $x < x$.

Сечение множества действительных чисел, образованное множествами A и B, обозначается через $A|B$. Множество A называется *нижним*, а множество B — *верхним* классом данного сечения.

Простые примеры сечений можно получить следующим образом. Зафиксируем какое-либо число $\alpha \in \boldsymbol{R}$. Отнесем сначала к множеству A все числа $x \leqslant \alpha$, а к множеству B — все числа $y > \alpha$:

$$A \overset{\text{def}}{=} \{x : x \leqslant \alpha\}, \quad B \overset{\text{def}}{=} \{y : y > \alpha\}. \tag{2.10}$$

Так определенные множества A и B образуют сечение, что устанавливается непосредственной проверкой выполнения условий 1^0), 2^0) и 3^0) определения 3.

Можно поступить иначе: отнести к множеству A все числа $x < \alpha$, а к множеству B — все числа $y \geqslant \alpha$:

$$A \overset{\text{def}}{=} \{x : x < \alpha\}, \quad B \overset{\text{def}}{=} \{y : y \geqslant \alpha\}. \tag{2.11}$$

Снова множества A и B образуют сечение. В обоих случаях (2.10) и (2.11) говорят, что сечение производится числом α, и пишут $\alpha = A|B$.

Отметим два свойства сечений, производящихся некоторым числом.

1^0. *В случае* (2.10) *в классе A есть наибольшее число, им является число* α, *а в классе B нет наименьшего числа.*

В случае (2.11) *в классе A нет наибольшего, а в классе B есть наименьшее число, им является число* α.

Рассмотрим, например, первый случай (2.10). То, что α является наибольшим числом в классе A, ясно из первой формулы (2.10), задающей множество A.

Покажем, что в множестве B нет наименьшего числа. Допустим противное: пусть в B есть наименьшее число. Обозначим его через β. Из условия $\beta \in B$, в силу второй формулы (2.10), справедливы неравенства $\alpha < \beta$, следовательно, $\alpha + \alpha < \alpha + \beta$, т. е. $\alpha < \frac{\alpha + \beta}{2}$, откуда снова, в силу второй формулы (2.10), получаем, что $\frac{\alpha + \beta}{2} \in B$. Аналогично из $\alpha < \beta$ имеем $\alpha + \beta < \beta + \beta$, т. е. $\frac{\alpha + \beta}{2} < \beta$, и так как β — наименьшее число в классе B, то $\frac{\alpha + \beta}{2} \in A$. Полученное противоречие доказывает утверждение. □

2^0. *Число, производящее сечение, единственно.*

В самом деле, допустим, что существует сечение, которое определяется двумя разными числами: $\alpha = A|B$ и $\beta = A|B$. Пусть, например, $\alpha < \beta$. Тогда, как было показано при доказательстве предыдущего свойства, справедливы неравенства $\alpha < \frac{\alpha + \beta}{2} < \beta$. Из неравенства $\alpha < \frac{\alpha + \beta}{2}$ следует, что как в случае (2.10), так и в случае (2.11) имеет место условие $\frac{\alpha + \beta}{2} \in B$. Аналогично из неравенства $\frac{\alpha + \beta}{2} < \beta$ следует, что $\frac{\alpha + \beta}{2} \in A$. Это противоречит тому, что множества A и B не пересекаются. □

Свойство непрерывности действительных чисел состоит в том, что никаких других сечений действительных чисел, кроме тех, которые производятся некоторым числом, не существует. Иначе говоря, непрерывность действительных чисел может быть описана следующим образом.

V_1. *Для каждого сечения $A|B$ множества действительных чисел существует число α, производящее это сечение*:

$$\alpha = A|B.$$

Это число, согласно доказанному выше, является либо наибольшим в нижнем классе, тогда в верхнем классе нет наименьшего числа, либо наименьшим в верхнем классе, тогда в нижнем нет наибольшего числа.

Таким образом, если $A|B$ является сечением множества действительных чисел, то, согласно свойству их непрерывности, сформулированному в форме V_1, не может случиться так, что в классе A будет наибольшее число и одновременно в классе B будет наименьшее число (рис. 4, *а*). Не может также быть и того, чтобы в классе A не было наибольшего числа и одновременно в классе B не было наименьшего числа (рис. 4, *б*). Образно говоря, непрерывность действительных чисел означает, что в их множестве нет ни скачков, ни пробелов, короче, нет пустот.

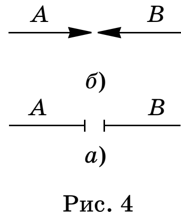


Рис. 4

Сечение $A|B$ геометрически означает разбиение числовой прямой на два луча, имеющих общее начало и идущих в противоположных направлениях, причем один из них содержит их общее начало (замкнутый луч), а другой — нет (открытый луч).

Сформулированное свойство непрерывности действительных чисел V, так же как и эквивалентное ему свойство V_1, называется *принципом непрерывности действительных чисел по Дедекинду*[1].

В дальнейшем мы встретимся с другими подходами к понятию непрерывности множества действительных чисел (см. п. 3.7).

Покажем, что свойство V_1 равносильно свойству V. Пусть сначала выполнено свойство V и задано какое-либо сечение $A|B$. В силу требования 3^0 в определении сечений для любых $a \in A$ и $b \in B$ выполняется неравенство $a < b$, поэтому пара множеств A и B удовлетворяет условию, сформулированному в свойстве V. Следовательно, согласно этому свойству, существует такое число α, что для всех $a \in A$ и $b \in B$ выполняется соотношение $a \leqslant \alpha \leqslant b$. Число α, согласно первому свойству сечений, принадлежит одному из классов A или B. Если $\alpha \in A$, то для всех $a \in A$ и $b \in B$ выполняется неравенство $a \leqslant \alpha < b$, т. е. число α производит сечение $A|B$ и является наибольшим в нижнем классе. Аналогично, если $\alpha \in B$, то число α также производит сечение $A|B$ и является наименьшим в верхнем классе B.

[1] Р. Дедекинд (1831—1916) — немецкий математик.

Пусть теперь, наоборот, выполнено условие V_1 и заданы два таких непустых множества $A \subset \boldsymbol{R}$ и $B \subset \boldsymbol{R}$, что для любых $a \in A$ и $b \subset \boldsymbol{R}$ выполняется неравенство $a \leqslant b$. Обозначим через B^* такое множество чисел, что, каково бы ни было $b^* \in B^*$, для любого $a \in A$ выполняется неравенство $a \leqslant b^*$ (число b^*, обладающее этим свойством, называется *числом, ограничивающим сверху множество* A). Очевидно,

$$B \subset B^*. \tag{2.12}$$

Через A^* обозначим все остальные действительные числа:

$$A^* = \boldsymbol{R} \backslash B^*.$$

Покажем, что множества A^* и B^* образуют сечение в множестве действительных чисел и что число α, производящее это сечение, удовлетворяет условию, указанному в формулировке свойства V, для заданных множеств A и B.

Прежде всего проверим, что множества A^* и B^* удовлетворяют всем условиям, которым должны удовлетворять множества, образующие сечение. Действительно, в множество A^* отнесены все числа, не попавшие в множество B^*, поэтому их объединение $A^* \cup B^*$ является множеством всех действительных чисел $\boldsymbol{R}$:

$$A^* \cup B^* = \boldsymbol{R}. \tag{2.13}$$

Множество B^* заведомо не пусто в силу включения (2.12), так как, по условию, множество B не пусто. Итак,

$$B^* \neq \varnothing. \tag{2.14}$$

Докажем, что и множество A^* не пусто. По условию, множество A не пусто[1]. Это означает, что существует по крайней мере одно число $a \in A$. Тогда число $a - 1$ заведомо не принадлежит множеству B^*. Действительно, $a - 1 < a$, а все числа b^* из множества B^* удовлетворяют неравенству $b^* \geqslant a$. По-

[1] Заметим, что не обязательно $A \subset A^*$. Более того, в случае, когда множество A состоит из одной точки $a \in B$ (это допустимо), множества A и A^* даже не пересекаются, так как в этом случае $A = \{a\} \subset B \subset B^*$.

этому $a - 1 \in A^*$, поскольку к множеству A^* относятся все числа, не вошедшие в множество B^*. Итак, множество A^* также не пусто:

x

α a_0

Рис. 5

$$A^* \neq \varnothing. \qquad (2.15)$$

Докажем теперь, что каждое число $a^* \in A^*$ меньше любого числа $b^* \in B^*$:

$$a^* < b^*. \qquad (2.16)$$

Допустим противное: пусть найдутся такие числа $a^* \in A^*$ и $b^* \in B^*$, что $a^* \geqslant b^*$. Тогда, в силу определения множества B^*, для любого $a \in A$ выполняется неравенство $a \leqslant b^*$, а следовательно, и неравенство $a \leqslant a^*$. Это означает, что $a^* \in B^*$. Таким образом, число a^* одновременно принадлежит как множеству A^*, так и множеству B^*. Это невозможно, так как к множеству A^* были отнесены только те числа, которые не содержатся в множестве B^*. Полученное противоречие показывает, что неравенство $a^* \geqslant b^*$ при условии $a^* \in A$, $b^* \in B$ невозможно, а тем самым выполняется неравенство (2.16).

Выполнение условий (2.13)—(2.16) означает, что множества A^* и B^* действительно образуют сечение в множестве действительных чисел (см. определение 3).

Пусть α — число, производящее это сечение: $\alpha = A^*|B^*$. Такое число α существует в силу предположения о выполнении свойства V_1. Покажем, что $\alpha \in B^*$. Если бы это было не так, то нашлось бы такое число $a_0 \in A$, что $a_0 > \alpha$. Выберем какое-либо число x так, чтобы $\alpha < x < a_0$ (рис. 5), например, $\frac{\alpha + a_0}{2}$. Из условий $x > \alpha$ и $\alpha = A^*|B^*$ вытекает, что $x \in B^*$, следовательно, для любого $a \in A$ должно выполняться неравенство $x \geqslant a$, так как B^* состоит только из таких чисел. Однако это неравенство не выполняется при $a = a_0$. Полученное противоречие доказывает, что $\alpha \in B^*$, и поэтому число α является наименьшим в верхнем классе B^*; но $B \subset B^*$, следо-

вательно, для любых $b \in B$ выполняется неравенство $\alpha \leqslant b$. Наконец, в силу самого определения множества B^*, из включения $\alpha \in B^*$ вытекает, что для любого числа $a \in A$ справедливо неравенство $a \leqslant \alpha$.

Итак, для всех $a \in A$, $b \in B$ имеет место неравенство $a \leqslant \alpha \leqslant$ $\leqslant b$. Это и означает, что наличие свойства V^* влечет за собой наличие свойства V. □

Аналогично определению 3 формулируется определение сечения в множестве рациональных чисел $\boldsymbol{Q}$. Всякое действительное число производит такое сечение, и можно доказать, что всякое сечение в $\boldsymbol{Q}$ производится действительным числом.

УПРАЖНЕНИЕ 1. Пусть $B = \{x : x^2 > 2,\ x > 0,\ x \in \boldsymbol{Q}\}$ и $A = \boldsymbol{Q} \backslash B$. Доказать, что множества A и B образуют сечение в поле рациональных чисел $\boldsymbol{Q}$ и что это сечение не определяется никаким рациональным числом.

2.6*. Рациональные степени действительных чисел

Число b такое, что $b^n = a$ (если оно, конечно, существует), называется *корнем* n-й степени из числа a и обозначается через $\sqrt[n]{a}$, или $a^{1/n}$, т. е. $(\sqrt[n]{a})^n \overset{\text{def}}{=} a$.

Отметим, что в множестве действительных чисел для любого числа $a \geqslant 0$ и любого натурального числа n всегда существует число $b \geqslant 0$, являющееся корнем n-й степени из a, т. е. существует $\sqrt[n]{a}$. Мы не будем пока останавливаться на доказательстве этого важного утверждения, хотя его можно было бы провести и здесь, например на основе понятия сечения, а докажем его позже (см. пример в п. 7.2). Конечно, в некоторых случаях корень может существовать и для $a < 0$. Например, существует $\sqrt[3]{-8} = -2$, но уже корень $\sqrt{-4}$ не существует в том смысле, что не существует действительного числа $b = \sqrt{-4}$, так как в противном случае было бы справедливо равенство $b^2 =$ $= -4$, которое противоречит правилу знаков при умножении.

Если $a \geqslant 0$, $b = \sqrt[n]{a}$ и $b \geqslant 0$, то число b называется *арифметическим значением* корня n-й степени из числа a. В дальнейшем под корнем из неотрицательного действительного числа будем понимать арифметическое значение корня, если не оговорено что-либо другое.

Сформулируем **свойства корня**. Пусть n и m — натуральные числа и $a \geqslant 0$, $b \geqslant 0$; тогда справедливы следующие формулы:

1^0. $\sqrt[n]{\sqrt[m]{a}} = \sqrt[mn]{a}$.

2^0. $\sqrt[n]{a} = \sqrt[mn]{a^m}$.

3^0. $\sqrt[n]{ab} = \sqrt[n]{a}\,\sqrt[n]{b}$.

4^0. $\sqrt[n]{\frac{a}{b}} = \frac{\sqrt[n]{a}}{\sqrt[n]{b}}$, $b \neq 0$.

$$5^0.\ (\sqrt[n]{a})^m = \sqrt[n]{a^m}.$$

Все эти формулы доказываются одинаковым приемом. Докажем, например, первую из них.

Пусть $b = \sqrt[n]{\sqrt[m]{a}}$. Согласно определению корня и свойству 22^0 из п. 2.2*, это означает, что $b^n = \sqrt[m]{a}$ и что $b^{mn} = a$. Отсюда, в силу того же определения корня, следует, что $b = \sqrt[mn]{a}$. Таким образом, имеем $\sqrt[n]{\sqrt[m]{a}} = b = \sqrt[mn]{a}$. □

Если $a < 0$ и все корни, входящие в формулу 1^0, существуют, то она также справедлива и приведенное ее доказательство сохраняет силу. Вообще, если $a < 0$ и все корни, входящие в какую-либо из формул 1^0—5^0, существуют, то они справедливы при соответствующем выборе значений корней.

Имея понятие целочисленной степени и корня, определим понятие рациональной степени. Пусть $a > 0$ и $r \in \boldsymbol{Q}$, т. е. $r = m/n$, $m \in \boldsymbol{Z}$, $n \in \boldsymbol{Z}$, $n \neq 0$.

Степень a^r определяется равенством $a^r \stackrel{\text{def}}{=} \sqrt[n]{a^m}$.

Отметим **основные свойства рациональной степени**. Пусть $a > 0$, $b > 0$, $r_1 \in \boldsymbol{Q}$, $r_2 \in \boldsymbol{Q}$; тогда:

6^0. $a^{-r} = \frac{1}{a^r}$.

7^0. $a^{r_1}a^{r_2} = a^{r_1 + r_2}$.

8^0. $(a^{r_1})^{r_2} = a^{r_1 r_2}$.

9^0. $(ab)^r = a^r b^r$.

Докажем свойство 6^0. Если $r = \frac{m}{n}$, где m и n — целые числа, $n \neq 0$, то

$$a^{-r} = a^{-\frac{m}{n}} = a^{\frac{-m}{n}} = \sqrt[n]{a^{-m}} = \sqrt[n]{\frac{1}{a^m}} = \frac{1}{\sqrt[n]{a^m}} = \frac{1}{a^{\frac{m}{n}}} = \frac{1}{a^r}. \square$$

Из формул 7^0, 8^0 и 9^0 докажем, например, формулу 7^0 (остальные доказываются аналогично). Если $r_1 = \frac{p}{q}$, $r_2 = \frac{m}{n}$, $q \neq 0$, $n \neq 0$, p, q, m, $n \in \boldsymbol{Z}$, то, использовав определение рациональной степени, свойства корней 2^0, 3^0 и свойство 22^0 из 2.2*, получим

$$a^{r_1}a^{r_2} = a^{\frac{p}{q}}a^{\frac{m}{n}} = \sqrt[q]{a^p}\sqrt[n]{a^m} = \sqrt[nq]{a^{np}}\sqrt[nq]{a^{mq}} =$$

$$= \sqrt[nq]{a^{np + mq}} = a^{\frac{np + mq}{nq}} = a^{\frac{p}{q} + \frac{m}{n}} = a^{r_1 + r_2}. \square$$

Из свойства 9^0 следует, что $\left(\frac{a}{b}\right)^r = \frac{a^r}{b^r}$. Действительно,

$$\left(\frac{a}{b}\right)^r = (ab^{-1})^r = a^r(b^{-1})^r = a^r b^{-r} = \frac{a^r}{b^r}. \square$$

Задача 1. Доказать с помощью сечений, что для любого числа $a > 0$ и любого натурального n существует корень $\sqrt[n]{a}$.

2.7. Формула бинома Ньютона

Многочлены, представляющие собой сумму двух слагаемых, называются двучленами или биномами. Выведем формулу для n-й степени бинома $x + a$, называемую *формулой бинома Ньютона*[1]:

$$(x + a)^n = x^n + C_n^1 x^{n-1}a + C_n^2 x^{n-2}a^2 + \ldots + C_n^{n-1} xa^{n-1} + a^n. \quad (2.17)$$

С помощью символа Σ формула (2.17) записывается в виде

$$(x + a)^n = \sum_{k=0}^{n} C_n^k x^{n-k}a^k. \quad (2.18)$$

Приведем два доказательства этой формулы: первое — эвристическое, оно позволит одновременно установить вид

[1] И. Ньютон (1643—1727) — английский физик, механик, математик и теолог.

формулы (2.18), второе, более короткое, будет состоять просто в проверке написанной формулы (2.18).

Первое доказательство. Вычислим произведение n биномов, вторые члены которых обозначены буквой a с различными индексами:

$$(x + a_1)(x + a_2) \dots (x + a_n). \tag{2.19}$$

Очевидно, что в произведении получится многочлен степени n относительно x. Сосчитаем последовательно коэффициенты у различных степеней x, начиная с наибольшей.

В произведении $(x + a_1)$, $(x + a_2)$, ... , $(x + a_n)$ член с x^n может получиться только в результате перемножения x-в, взятых из каждой скобки; поэтому коэффициент у x^n равен 1:

$$(x + a_1)(x + a_2) \dots (x + a_n) = x^n + \dots .$$

Член с x^{n-1} при вычислении произведения (2.19) может получиться только в том случае, когда во всех скобках, кроме одной, выбран x, а из оставшейся — свободный член, которым может быть любое из a_1, a_2, ... , a_n. Поэтому коэффициент в произведении (2.19) при x^{n-1} равен $a_1 + a_2 + \dots a_n$:

$$(x + a_1)(x + a_2) \dots (x + a_n) = x^n + x^{n-1}(a_1 + a_2 + \dots + a_n) + \dots .$$

Член с x^{n-2} может получиться только если из всех сомножителей произведения (2.19), кроме двух, выбран член x, а в оставшихся двух — свободные члены; поэтому коэффициент при x^{n-2} в произведении (2.19) является суммой всевозможных произведений $a_{i_1}a_{i_2}$ различных a_{i_1} и a_{i_2}, причем их всегда можно расположить в порядке возрастания индексов, т. е. так, что $i_1 < i_2$, $i_1, i_2 = 1, 2, \dots, n$:

$$(x + a_1)(x + a_2) \dots (x + a_n) =$$
$$= x^n + x^{n-1}\sum_{i=1}^{n} a_i + x^{n-2}\sum_{\substack{i_1, i_2 = 1 \\ i_1 < i_2}}^{n} a_{i_1}a_{i_2} + \dots .$$

Число слагаемых в сумме $\sum\limits_{\substack{i_1, i_2 = 1 \\ i_1 < i_2}}^{n} a_{i_1}a_{i_2}$ равно числу всевозможных выборов пар различных членов из a_1, a_2, ... , a_n, т. е. числу сочетаний C_n^2.

Вообще, член с x^{n-k}, $k = 1, 2, \dots, n-1$, может получиться только если из всех сомножителей произведения (2.19), кроме k, выбран член x, а в оставшихся k — свободные члены. Поэтому коэффициент при x^{n-k} в произведении (2.19) является суммой всевозможных произведений $a_{i_1}a_{i_2}\dots a_{i_k}$ различных a_{i_j}. Сомножители a_{i_j} в этих произведениях всегда можно расположить в порядке возрастания их индексов, т. е. считать, что $i_1 < i_2 < \dots < i_k$, $i_j \in \{1, 2, \dots, n\}$, $j = 1, 2, \dots, k$:

$$(x + a_1)(x + a_2) \dots (x + a_n) = x^n + x^{n-1}\sum_{k=1}^{n} a_i +$$

$$+ x^{n-2} \sum_{\substack{i_1,\, i_2 = 1 \\ i_1 < i_2}}^{n} a_{i_1}a_{i_2} + \dots + x^{n-k} \sum_{\substack{i_1,\, i_2,\, \dots,\, i_k = 1 \\ i_1 < i_2 < \dots < i_k}}^{n} a_{i_1}a_{i_2}\dots a_{i_k} + \dots .$$

Число слагаемых в сумме, являющейся коэффициентом при x^{n-k}, равно числу сочетаний из n элементов по k, т. е. равно C_n^k.

Наконец, свободный член произведения (2.19) получится лишь в том случае, когда перемножаются свободные члены каждой из скобок, т. е. свободный член произведения (2.19) равен $a_1a_2\dots a_n$. Итак,

$$(x + a_1)(x + a_2) \dots (x + a_n) = x^n + x^{n-1}\sum_{i=1}^{n} a_i + x^{n-2}\sum_{\substack{i_1,\, i_2 = 1 \\ i_1 < i_2}}^{n} a_{i_1}a_{i_2} +$$

$$+ \dots + x^{n-k} \sum_{\substack{i_1,\, i_2,\, \dots,\, i_k = 1 \\ i_1 < i_2 < \dots < i_k}}^{n} a_{i_1}a_{i_2}\dots a_{i_k} + \dots + a_1a_2\dots a_n. \qquad (2.20)$$

Число слагаемых в каждой из сумм, написанных со знаком Σ, равно, соответственно, C_n^1, C_n^2, ..., C_n^k, ... C_n^{n-1}, а у первого и последнего слагаемого для записи коэффициента 1 можно использовать равенства $C_n^0 = C_n^n = 1$. Поэтому при $a_1 = a_2 = \dots = a_n = a$ из формулы (2.20) получим

$$(x + a)^n = C_n^0 x^n + C_n^1 x^{n-1}a +$$
$$+ C_n^2 x^{n-2}a^2 + \dots + C_n^k x^{n-k}a^k + \dots + C_n^n a^n,$$

т. е. формулу (2.18).

Второе доказательство проведем методом математической индукции (для него надо заранее знать формулу (2.18), тогда как и при первом доказательстве мы ее нашли).

При $n = 1$ формула (2.18) очевидно верна, так как в этом случае она имеет вид $x + a = x + a$.

Пусть формула (2.18) верна при некотором $n \in \boldsymbol{N}$. Докажем, что тогда она верна и при $n + 1$:

$$(x + a)^{n+1} = (x + a)^n(x + a) \underset{(2.17)}{=} (x^n + C_n^1 x^{n-1}a + C_n^2 x^{n-2}a^2 + \ldots$$

$$\ldots + C_n^k x^{n-k}a^k + C_n^{k+1} x^{n-k-1}a^{k+1} + \ldots + a^n)(x + a) =$$

$$= x^{n+1} + C_n^1 x^n a + C_n^2 x^{n-1}a^2 + \ldots + C_n^{k+1} x^{n-k}a^{k+1} + \ldots + xa^n +$$

$$+ x^n a + C_n^1 x^{n-1}a^2 + \ldots + C_n^k x^{n-k} a^{k+1} + \ldots + a^{n+1} =$$

$$= x^{n+1} + (C_n^1 + C_n^0)x^n a + (C_n^2 + C_n^1)x^{n-1}a^2 + \ldots$$

$$\ldots + (C_n^{k+1} + C_n^k)x^{n-k}a^{k+1} + \ldots + a^{n+1}.$$

Вспомнив, что $C_n^{k+1} + C_n^k = C_{n+1}^{k+1}$ (см. (1.19)), получим

$$(x + a)^{n+1} = x^{n+1} + C_{n+1}^1 x^n a +$$

$$+ C_{n+1}^2 x^{n+1}a^2 + \ldots + C_{n+1}^{k+1} x^{n-k}a^{k+1} + \ldots + a^{n+1},$$

т. е. формулу (2.17), в которой n заменено на $n + 1$. □

З а м е ч а н и е. Отметим два интересных свойства биномиальных коэффициентов C_n^k, $k = 0, 1, \ldots, n$: их сумма равна 2^n, причем сумма коэффициентов, стоящих на четных местах, равна сумме коэффициентов, стоящих на нечетных местах. Действительно, подставив в формулу бинома Ньютона (2.18) $x = a = 1$, получим $\sum_{k=0}^{n} C_n^k = 2^n$, а подставив $x = 1$, $a = -1$, имеем $\sum_{k=0}^{n} (-1)^k C_n^k = 0$.

УПРАЖНЕНИЕ 2. Доказать, что

$$(x_1 + x_2 + \ldots + x_n)^m = \sum_{m_1 + \ldots + m_n = 0}^{n} \frac{m!}{m_1! \ldots m_n!} x_1^{m_1} \ldots x_n^{m_n}.$$

§ 3.

Числовые множества

3.1. Расширенная числовая прямая

Часто бывает удобно дополнить множество действительных чисел $\boldsymbol{R}$ элементами, обозначаемыми через $+\infty$ и $-\infty$ и назы-

ваемыми соответственно плюс и минус бесконечностями, считая при этом, что, по определению,

$$-\infty < +\infty,$$
$$(+\infty) + (+\infty) = +\infty,\ (-\infty) + (-\infty) = -\infty,$$
$$+\infty - (-\infty) = +\infty,\ -\infty - (+\infty) = -\infty,$$
$$(+\infty)(+\infty) = (-\infty)(-\infty) = +\infty,$$
$$(+\infty)(-\infty) = (-\infty)(+\infty) = -\infty.$$

Но, например, операции $(+\infty) + (-\infty)$ или $\frac{+\infty}{+\infty}$ уже не определены. Кроме того, для любого $a \in \boldsymbol{R}$ по определению полагается выполненным неравенство $-\infty < a < +\infty$ и справедливость следующих операций:

$$a + (+\infty) = +\infty + a = +\infty,\ -\infty + a = a + (-\infty) = -\infty;$$

для $a > 0$

$$a(+\infty) = (+\infty)a = +\infty,\ a(-\infty) = (-\infty)a = -\infty;$$

для $a < 0$

$$a(+\infty) = (+\infty)a = -\infty,\ a(-\infty) = (-\infty)a = +\infty.$$

Бесконечности $+\infty$ и $-\infty$ называют иногда «*бесконечными числами*» в отличие от действительных чисел $a \in \boldsymbol{R}$, которые называются также *конечными числами*.

Множество действительных чисел $\boldsymbol{R}$, дополненное элементами $+\infty$ и $-\infty$, называется *расширенным множеством действительных* чисел (или расширенной числовой прямой) и обозначается через $\overline{\boldsymbol{R}}$. Элементы $+\infty$ и $-\infty$ называются иногда бесконечно удаленными точками расширенной числовой прямой в противопоставление точкам числовой прямой $\boldsymbol{R}$, которые называются также и *конечными точками*.

В дальнейшем под числом всегда понимается конечное действительное число, если не оговорено что-либо другое.

3.2. Промежутки действительных чисел. Окрестности

Напомним определения некоторых основных подмножеств действительных чисел, которые часто будут встречаться в дальнейшем. Если $a \leqslant b$, $a \in \boldsymbol{R}$, $b \in \boldsymbol{R}$, то множество $\{x : a \leqslant x \leqslant b\}$

называется *отрезком* расширенной числовой прямой $\boldsymbol{R}$ и обозначается через $[a,\ b]$, т. е.

$$[a,\ b] \stackrel{\text{def}}{=} \{x : a \leqslant x \leqslant b\},\ a \in \overline{\boldsymbol{R}},\ b \in \overline{\boldsymbol{R}}.$$

В случае $a = b$ отрезок $[a,\ b]$ состоит из одной точки.

Если $a < b$, то множество $\{x : a < x < b\}$ называется *интервалом* и обозначается через $(a,\ b)$, т. е.

$$(a,\ b) \stackrel{\text{def}}{=} \{x : a < x < b\}.$$

Интервал $(a,\ b)$ называется *внутренностью отрезка* $[a,\ b]$.

Множества

$$[a,\ b) \stackrel{\text{def}}{=} \{x : a \leqslant x < b\} \text{ и } (a,\ b] \stackrel{\text{def}}{=} \{x : a < x \leqslant b\}$$

называются *полуинтервалами*.

Отрезки $[a,\ b]$, интервалы $(a,\ b)$ и полуинтервалы $[a,\ b)$, $(a,\ b]$ называются *промежутками*, точки a и b — их *концами*: a — левым концом, а b — правым, а точки x такие, что $a < x < b$, — их *внутренними точками*.

Если a и b конечны, т. е. $a \in \boldsymbol{R}$ и $b \in \boldsymbol{R}$, то промежуток с концами a и b называется также *конечным промежутком*, а число $b - a$ — его *длиной*.

Если хотя бы одно из a и b является бесконечным, то промежуток с концами a и b называется *бесконечным*.

Замечание 1. Промежутки всех типов расширенной числовой прямой обладают следующим свойством: *если точки* $\alpha \in \overline{\boldsymbol{R}}$ *и* $\beta \in \overline{\boldsymbol{R}}$, $\alpha < \beta$, *принадлежат некоторому промежутку с концами* $a \in \overline{\boldsymbol{R}}$ *и* $b \in \overline{\boldsymbol{R}}$, *то и весь отрезок* $[\alpha,\ \beta]$ *принадлежит этому промежутку.*

Для промежутка каждого типа это непосредственно следует из его определения.

Важным понятием для дальнейшего является понятие ε-окрестности точки расширенной числовой прямой.

В случае $a \in \boldsymbol{R}$, т. е. когда a является действительным числом, *ε-окрестностью* $U(a,\ \varepsilon)$[1], $\varepsilon > 0$, *числа* a называется интервал $(a - \varepsilon,\ a + \varepsilon)$:

$$U(a,\ \varepsilon) \stackrel{\text{def}}{=} (a - \varepsilon,\ a + \varepsilon).$$

---

[1] Обозначение окрестности точки символом U происходит от слова Ungebung (нем.) — окрестность.

Если же $a = +\infty$, то

$$U(+\infty, \varepsilon) \stackrel{\text{def}}{=} \left(\frac{1}{\varepsilon}, +\infty\right],$$

а если $a = -\infty$, то

$$U(-\infty, \varepsilon) \stackrel{\text{def}}{=} \left[-\infty, -\frac{1}{\varepsilon}\right).$$

Таким образом, во всех случаях, т. е. когда a — действительное число или когда a — одна из бесконечностей $+\infty$, $-\infty$, при уменьшении числа ε соответствующие ε-окрестности $U(a, \varepsilon)$ уменьшаются: если $0 < \varepsilon_1 < \varepsilon_2$, то $U(a, \varepsilon_1) \subset U(a, \varepsilon_2)$.

Иногда бывает удобно пополнить множество действительных чисел не двумя, а одной бесконечностью (без знака) ∞. Ее ε-окрестность $U(\infty, \varepsilon)$, $\varepsilon > 0$, определяется равенством

$$U(\infty, \varepsilon) \stackrel{\text{def}}{=} \left\{x : x \in \boldsymbol{R}, |x| > \frac{1}{\varepsilon}\right\} \cup \{\infty\}.$$

Иначе говоря, ε-окрестность $U(\infty, \varepsilon)$ состоит из двух бесконечных интервалов $\left(-\infty, -\frac{1}{\varepsilon}\right)$, $\left(\frac{1}{\varepsilon}, +\infty\right)$ и самого элемента ∞. Этот элемент также называется иногда бесконечно удаленной точкой числовой прямой. В отличие от бесконечностей со знаком $+\infty$ и $-\infty$ бесконечность ∞ без знака не связана с действительными числами отношением порядка.

Всякая ε-окрестность конечной или бесконечно удаленной точки числовой прямой называется ее *окрестностью* и часто обозначается просто через $U(a)$. Иногда мы будем обозначать окрестности и другими буквами, например V, W.

Нередко бывает удобным называть окрестностью $U(a)$ конечной точки a не только интервал, центром которого является точка a, а всякий ее содержащий интеграл. Ясно, что любая такая окрестность точки содержит ее ε-окрестность при некотором $\varepsilon > 0$.

Наряду с определенными выше окрестностями бесконечностей в пополнениях ими множества действительных чисел иногда рассматривают и окрестности бесконечностей ∞, $+\infty$ и $-\infty$ в самом множестве действительных чисел: $U(\infty) \cap \boldsymbol{R}$, $U(+\infty) \cap \boldsymbol{R}$ и $U(-\infty) \cap \boldsymbol{R}$. Сами бесконечности, конечно, уже не попадают в эти окрестности. Мы будем придерживаться первоначально данных определений (отметим, впрочем, что доказываемая ниже лемма справедлива и в случае, когда в ней под окрестностями бесконечностей понимают их окрестности в множестве действительных чисел).

Сформулируем в виде леммы одно важное свойство окрестностей.

ЛЕММА. *У любых двух различных точек расширенной числовой прямой (расширенной с помощью двух бесконечностей со знаком или при помощи только одной бесконечности без знака) существуют непересекающиеся окрестности.*

Доказательство. Рассмотрим сначала случай расширенной числовой прямой $\overline{\boldsymbol{R}}$, полученной добавлением к множеству действительных чисел $\boldsymbol{R}$ двух бесконечностей со знаком. Покажем, что для любых $a \in \overline{\boldsymbol{R}}$ и $b \in \overline{\boldsymbol{R}}$, $a < b$, существуют такие $\varepsilon_1 > 0$ и $\varepsilon_2 > 0$, что $U(a, \varepsilon_1) \cap U(b, \varepsilon_2) = \varnothing$. В самом деле, если a и b — действительные числа, то можно взять $\varepsilon_1 = \varepsilon_2 = \frac{b-a}{2}$ (рис. 6, *а*). Если a — действительное число, а $b = +\infty$, то в качестве указанных $\varepsilon_1 > 0$ и $\varepsilon_2 > 0$ подходят, например, $\varepsilon_1 = 1$ и $\varepsilon_2 = \frac{1}{|a|+1}$ (рис. 6, *б*). Если $a = -\infty$, b — действительное число, то можно взять $\varepsilon_1 = \frac{1}{|b|+1}$, $\varepsilon_2 = 1$ (рис. 6, *в*). Наконец, если $a = -\infty$, $b = +\infty$, то при произвольном $\varepsilon > 0$ окрестности $U(-\infty, \varepsilon)$ и $U(+\infty, \varepsilon)$ не пересекаются (рис. 6, *г*). Если же числовая прямая $\boldsymbol{R}$ дополнена лишь одной бесконечностью ∞, то достаточно рассмотреть лишь случай $a \in \boldsymbol{R}$ и $b = \infty$ (так как случай $a \in \boldsymbol{R}$ и $b \in \boldsymbol{R}$ рассмотрен выше), в котором можно снова (как и при $a \in \boldsymbol{R}$, $b = +\infty$) взять $\varepsilon_1 = 1$, а $\varepsilon_2 = \frac{1}{|a|+1}$. □

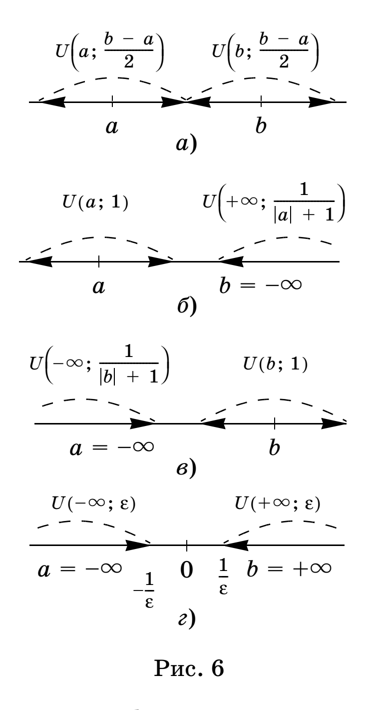


Рис. 6

Замечание 2. В случае $a < b$, $a \in \overline{\boldsymbol{R}}$, $b \in \overline{\boldsymbol{R}}$ и их непересекающихся окрестностей $U(a, \varepsilon_1) \cap U(b, \varepsilon_2) = \varnothing$ для любых $x \in U(a, \varepsilon_1)$ и $y \in U(b, \varepsilon_2)$, очевидно, справедливо неравенство $x < y$.

Его справедливость устанавливается непосредственной проверкой во всех возможных здесь случаях, т. е. при $a \in \boldsymbol{R}$, $b \in \boldsymbol{R}$, при $a \in \boldsymbol{R}$, $b = +\infty$, при $a = -\infty$, $b \in \boldsymbol{R}$ и при $a = -\infty$, $b = +\infty$.

Легко убедиться, что пересечение двух окрестностей точки (конечной или бесконечно удаленной) является также окрестностью этой точки.

3.3. Ограниченные и неограниченные множества

Введем ряд нужных для дальнейшего понятий и изучим некоторые свойства числовых множеств.

Определение 1. *Если для подмножества X действительных чисел существует такое число b, что оно не меньше каждого числа $x \in X$, т. е. для любого $x \in X$ выполняется неравенство $x \leqslant b$, то множество X называется ограниченным сверху, а число b — числом, ограничивающим сверху множество X.*

С помощью логических символов определение ограниченного сверху множества записывается в виде:

X ограничено сверху $\Leftrightarrow \exists\, b \in \boldsymbol{R}\ \forall\, x \in X : x \leqslant b$;

отсюда

X не ограничено сверху $\Leftrightarrow \forall\, b \in \boldsymbol{R}\ \exists\, x \in X : x > b$,

т. е. множество X не ограничено сверху, если, каково бы ни было число $b \in \boldsymbol{R}$, найдется такое число $x \in X$, что $x > b$.

Множество, не являющееся ограниченным сверху множеством, называется *неограниченным сверху множеством.*

Заметим, что если число b ограничивает сверху множество X, т. е. для всех $x \in X$ выполняется неравенство $x \leqslant b$ и $b < b'$, то для всех $x \in X$, очевидно, имеет место и неравенство $x \leqslant b'$, следовательно, число b' также ограничивает сверху множество X.

Если в множестве X имеется число b, которое не меньше всех других чисел из X, т. е. $b \in X$, и для всех $x \in X$ выполняется неравенство $x \leqslant b$, то число b называется наибольшим или *максимальным числом множества* X: $b = \max X$.

Очевидно, что если в множестве X имеется наибольшее число, то оно единственно, а само множество X в этом случае ограничено сверху этим числом.

Отметим еще, что если множество X не ограничено сверху, то, согласно определению, это означает, что для любого числа $b \in \boldsymbol{R}$ существует по крайней мере один такой элемент $x \in X$, что $x > b$. Обратим внимание на то, что на самом деле таких элементов бесконечно много. Действительно, до-

пустим, что их оказалось лишь конечное число: $x_1, \ldots, x_n$, $n \in \boldsymbol{N}$. Иначе говоря, для всех $x \in X$ и $x \neq x_k$, $k = 1, 2, \ldots, n$, справедливо неравенство $x \leqslant b$. Тогда ясно, что для $b_0 = \max\{b, x_1, \ldots, x_n\}$ и всех $x \in X$ будет выполняться неравенство $x \leqslant b_0$, т. е. вопреки предположению множество X оказалось ограниченным.

Аналогично множеству, ограниченному сверху, определяется множество, ограниченное снизу.

Определение 2. *Если для подмножества X действительных чисел существует такое число a, что оно не больше каждого числа $x \in X$, т. е. для любого $x \in X$ выполняется неравенство $a \leqslant x$, то множество X называется ограниченным снизу, а число a — числом, ограничивающим снизу это множество.*

Множество, не являющееся ограниченным снизу множеством, называется *неограниченным снизу множеством*.

С помощью логических символов определение ограниченного снизу множества записывается в виде: X ограничено снизу $\Leftrightarrow \exists\, a \in \boldsymbol{R}\ \forall\, x \in X : x \geqslant a$; отсюда X не ограничено снизу $\Leftrightarrow \forall\, a \in \boldsymbol{R}\ \exists\, x \in X : x < a$, т. е. множество X не ограничено снизу, если, каково бы ни было число $a \in \boldsymbol{R}$, найдется такой элемент $x \in X$, что $x < a$.

Очевидно, что если число a ограничивает снизу множество X, то и любое число $a' < a$ также ограничивает снизу это множество.

Если в множестве X имеется число a, которое не больше всех других чисел из X, т. е. $a \in X$, и для всех $x \in X$ выполняется неравенство $a \leqslant x$, то число a называется наименьшим или *минимальным числом множества* X : $a = \min X$.

Если в множестве X имеется наименьшее число, то оно единственно, а само множество X в этом случае ограничено снизу этим числом.

Определение 3. *Множество, ограниченное и сверху и снизу, называется ограниченным множеством.*

Другими словами, множество $X \subset \boldsymbol{R}$ называется ограниченным, если существуют такие числа a и b, что для любого $x \in X$ выполняется неравенство $a \leqslant x \leqslant b$.

Множество, не являющееся ограниченным, называется *неограниченным*. Очевидно, что неограниченное множество

может быть неограниченным и сверху и снизу или только сверху или снизу.

УПРАЖНЕНИЕ 1. Доказать, что множество $X \subset \boldsymbol{R}$ ограничено тогда и только тогда, когда существует такое число $a \geqslant 0$, что для всех $x \in X$ выполняется неравенство $|x| \leqslant a$.

Примерами ограниченных множеств являются отрезок $[1, 2]$, интервал $(0, 1)$, множество значений функции $\sin x$. Бесконечный интервал $(-5, +\infty)$, множество натуральных чисел 1, 2, 3, ... являются множествами, ограниченными снизу, но неограниченными сверху. Наконец, множество всех целых чисел, всех рациональных чисел суть множества, неограниченные как сверху, так и снизу.

Формальное обобщение понятий ограниченного сверху, ограниченного снизу и просто ограниченного множества на подмножества расширенного множества $\overline{\boldsymbol{R}}$ действительных чисел $\boldsymbol{R}$ (см. п. 2.5) не приводит к содержательным понятиям, так как все подмножества расширенного множества действительных чисел ограничены сверху символом $+\infty$ и снизу символом $-\infty$, а поэтому и просто ограничены в $\overline{\boldsymbol{R}}$. Однако понятие наибольшего (наименьшего) элемента множества является содержательным и в этом случае. Его определение формально совпадает с соответствующим определением для подмножеств не расширенного множества действительных чисел:

конечное или бесконечное число $c \in X \subset \overline{\boldsymbol{R}}$ называется наибольшим (*наименьшим*) *в множестве $X \subset \overline{\boldsymbol{R}}$, если для всех $x \in X$ выполняется неравенство $x \leqslant c$* (*соответственно $x \geqslant c$*).

В дальнейшем мы воспользуемся этим понятием.

3.4. Верхняя и нижняя грани числовых множеств

Среди всех чисел, ограничивающих сверху (снизу) данное множество, наименьшее (наибольшее) из них имеет специальное название.

Определение 4. *Наименьшее среди всех чисел, ограничивающих сверху множество $X \subset \boldsymbol{R}$, называется его верхней гранью и обозначается*[1] $\sup X$ *или* $\sup\limits_{x \in X} \{x\}$.

---

[1] Лат. supremum — наибольший.

Определение 5. *Наибольшее среди всех чисел, ограничивающих снизу множество* $X \subset \boldsymbol{R}$, *называется его нижней гранью и обозначается*[1] $\inf X$ *или* $\inf\limits_{x \in X} \{x\}$.

Иногда верхнюю (нижнюю) грань множества называют *точной верхней* (*нижней*) *гранью* этого множества.

Отметим, что в сделанных определениях не обсуждается вопрос о том, существует или нет наименьшее (наибольшее) число среди всех чисел, ограничивающих сверху (снизу) данное множество, — это будет сделано позже. Здесь же лишь говорится, что если такое число существует, то оно называется верхней (нижней) гранью рассматриваемого множества. Из самого определения верхней (нижней) грани следует, что если у данного множества эта грань существует, то она единственна, так как во всяком множестве максимальное (минимальное) число может быть только одно.

Проанализируем определения 4 и 5. Пусть $\beta = \sup X$. Это означает, во-первых, что число β ограничивает сверху множество X, т. е. для каждого $x \in X$ справедливо неравенство $x \leqslant \beta$; во-вторых, что число β является наименьшим среди всех чисел, ограничивающих сверху множество X, т. е., каково бы ни было число $\beta' < \beta$, оно уже не ограничивает сверху множество X, а это означает, что в множестве X найдется такое число x, что $x > \beta'$.

Таким образом, в «арифметической форме» определение 4 можно записать в следующем виде.

Определение 4′. *Число* β *называется верхней гранью множества* X, *если*:

1^0) $\forall\, x \in X : x \leqslant \beta$;

2^0) $\forall\, \beta' < \beta\ \exists\, x \in X : x > \beta'$.

Условие 2^0) можно переписать следующим образом:

$2'$) $\forall\, \varepsilon > 0\ \exists\, x \in X : x > \beta - \varepsilon$.

Для того чтобы убедиться в равносильности условий 2^0 и $2'$, достаточно взять β' и β, связанные равенством $\beta' = \beta - \varepsilon$, из которого следует, что условие $\varepsilon > 0$ эквивалентно условию $\beta' < \beta$.

---

[1] Лат. infimum — наименьший.

Аналогичным образом, если $\alpha = \inf X$, то, согласно определению 5, во-первых, число α ограничивает снизу множество X, а во-вторых, любое число $\alpha' > \alpha$ уже не ограничивает снизу это множество, ибо число α является наибольшим среди всех таких чисел. Это означает, что для любого $\alpha' > \alpha$ найдется такой $x \in X$, что $x < \alpha'$. Следовательно, определение 5 можно перефразировать следующим образом.

Определение 5′. *Число α называется нижней гранью множества X, если*:

1^0) $\forall\, x \in X : x \geqslant \alpha$;

2^0) $\forall\, \alpha' > \alpha\ \exists\, x \in X : x < \alpha'$.

Условие 2^0) эквивалентно условию

2′) $\forall\, \varepsilon > 0\ \exists\, x \in X : x < \alpha + \varepsilon$.

Для того чтобы убедиться в эквивалентности условий 2^0) и 2′), достаточно взять $\alpha' = \alpha + \varepsilon$.

Сделаем несколько очевидных замечаний. Если непустое множество $X \subset \boldsymbol{R}$ имеет верхнюю грань $\beta \in \boldsymbol{R}$ (имеет нижнюю грань $\alpha \in \boldsymbol{R}$), то оно ограничено сверху (снизу). Это следует из условия 1^0 определения 4′ (определения 5′).

Если $\beta = \sup X$ ($\alpha = \inf X$) и число b (число a) ограничивает сверху (снизу) множество X, то $\beta \leqslant b$ (соответственно $a \leqslant \alpha$). Это следует из того, что верхняя (нижняя) грань множества является наименьшим (наибольшим) числом среди всех чисел, ограничивающих сверху (снизу) данное множество.

Если в множестве существует наибольшее (наименьшее) число, то оно является верхней (нижней) гранью этого множества. В частности, такая ситуация имеет место для конечных множеств: любое конечное множество чисел имеет наибольшее и наименьшее числа, а потому нижнюю и верхнюю грани. В принципе их можно найти простым перебором всех чисел из данного множества, так как оно конечно. Однако, вообще говоря, только в принципе, а не на практике: если в данном конечном множестве, заданном какими-то свойствами его элементов, будет «достаточно много» элементов, то перебрать их все будет, возможно, не под силу даже сверхмощной современной вычислительной машине.

Приведем примеры, иллюстрирующие понятие верхней и нижней граней множества.

Множество всех положительных действительных чисел (обозначим его через $\boldsymbol{R}_+$) ограничено снизу числом нуль, ибо для любого $x \in \boldsymbol{R}_+$ имеет место $x > 0$; более того, $\inf \boldsymbol{R}_+ = 0$. Множество $\boldsymbol{R}_+$ не ограничено сверху, так как нет числа, которое бы ограничивало сверху все положительные числа.

Если $X = [a, b]$ — отрезок, то $\inf X = a$, $\sup X = b$. Если $X = (a, b)$ — интервал, то также $\inf X = a$, $\sup X = b$. Если, наконец, множество X состоит из двух точек a и b, $a \leqslant b$, т. е. $X = \{a\} \cup \{b\}$, то снова $\inf X = a$, $\sup X = b$. Эти примеры показывают, в частности, что верхняя (нижняя) грань множества может как принадлежать самому множеству, так и не принадлежать ему.

Если $\xi = A|B$ — сечение в области действительных чисел (см. п. 2.5*), то

$$\xi = \sup A = \inf B. \tag{3.1}$$

Выясним теперь вопрос: всегда ли у числового множества существует его верхняя (нижняя) грань? Если множество не ограничено сверху (снизу), то не существует чисел, которые бы ограничивали его сверху (снизу). Следовательно, не существует среди них и наименьшего (наибольшего). Таким образом, если множество не ограничено сверху (снизу), то у него нет верхней (нижней) грани. В этом случае ответ на поставленный вопрос получился совсем просто. Если же множество ограничено сверху (снизу), то ответ дается следующей теоремой.

ТЕОРЕМА 1. *Всякое ограниченное сверху непустое числовое множество имеет верхнюю грань, а всякое ограниченное снизу непустое числовое множество имеет нижнюю грань.*

Доказательство. Пусть X — ограниченное сверху непустое числовое множество. Обозначим через Y множество всех чисел, ограничивающих сверху множество X. Множество X ограничено сверху, поэтому множество Y не пусто. Каждый элемент $y \in Y$ ограничивает сверху множество X, т. е. для любого элемента $x \in X$ выполняется неравенство $x \leqslant y$. Элементы x и y являются произвольными элементами соответственно множеств X и Y, поэтому, в силу свойства непрерывности действительных чисел (см. свойство V в п. 2.1), существует такое число β, что для любых $x \in X$ и $y \in Y$ имеет место неравенство

$$x \leqslant \beta \leqslant y. \tag{3.2}$$

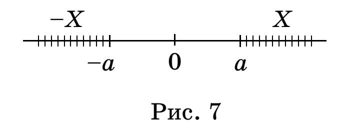


Рис. 7

Выполнение неравенства $x \leqslant \beta$ для всех $x \in X$ означает, что число β ограничивает сверху множество X, а выполнение неравенства $\beta \leqslant y$ для всех $y \in Y$, т. е. для всех чисел, ограничивающих сверху множество X, означает, что число β является наименьшим среди всех таких чисел, т. е. верхней гранью множества X:

$$\beta = \sup X. \tag{3.3}$$

Итак, существование верхней грани у ограниченного сверху непустого множества доказано.

Если теперь Y — непустое ограниченное снизу числовое множество, то отнесем к множеству X все числа, ограничивающие снизу множество Y. Далее, рассуждая аналогично рассмотренному случаю верхней грани, легко убеждаемся, что, в силу свойства непрерывности действительных чисел, существует такое число α, что для любых $x \in X$ и $y \in Y$ выполняется неравенство

$$x \leqslant \alpha \leqslant y. \tag{3.4}$$

Это, очевидно, и означает, что $\alpha = \inf Y$. □

Впрочем, утверждение о существовании нижней грани у ограниченного снизу непустого множества можно получить и из уже доказанного утверждения о существовании верхней грани у непустого ограниченного сверху множества. Достаточно заметить, что если X — ограниченное снизу множество, то множество $-X$ всех чисел $-x$, где $x \in X$, т. е. множество на числовой прямой, симметричное с множеством X относительно нуля, является уже ограниченным сверху множеством, и, наоборот, если X — ограниченное сверху множество, то множество $-X$ ограничено снизу (рис. 7). Действительно, если число a ограничивает снизу множество X, то число $-a$ ограничивает сверху множество $-X$, а если число b ограничивает сверху множество X, то число $-b$ ограничивает снизу множество $-X$. Отсюда следует, что

$$\sup(-X) = -\inf X, \quad \inf(-X) = -\sup X. \tag{3.5}$$

Из существования верхней грани у ограниченного сверху непустого множества и каждого из равенств (3.5) следует существование нижней грани у ограниченного снизу непустого множества.

Теорема о существовании верхних и нижних граней принадлежит к так называемым чистым теоремам существования: в ней доказывается, что при определенных условиях у множества существует верхняя (нижняя) грань. Однако из рассуждений, проведенных при доказательстве этой теоремы, не следует способ нахождения этих граней в конкретном случае. Это ясно из того, что построение множества Y, с помощью которого проводилось доказательство теоремы и которое состояло из всех чисел, ограничивающих сверху рассматриваемое множество, равносильно отысканию верхней грани β этого множества. В действительности задача нахождения верхней (нижней) грани множества, заданного какими-либо своими свойствами, может оказаться очень трудной задачей.

Если множество не ограничено сверху (снизу), то, как уже отмечалось, никакое число не может являться его верхней (нижней) гранью, так как вообще нет чисел, которые его ограничивают сверху (снизу). Для удобства вводится следующее определение.

Верхней гранью неограниченного сверху числового множества называется $+\infty$, *а нижней гранью неограниченного снизу числового множества называется* $-\infty$.

Это определение естественно, так как при соглашениях, принятых относительно употребления символов $+\infty$ и $-\infty$ в п. 2.5, определенные таким образом бесконечные грани множеств также удовлетворяют условиям 1^0 и 2^0 определений $4'$ и $5'$.

Удобство же этого определения состоит в том, что теперь *каждое непустое числовое множество имеет верхнюю грань, принадлежащую расширенному множеству действительных чисел*. При этом если заданное множество ограничено сверху, то его верхняя грань к о н е ч н а, если же оно не ограничено сверху, то б е с к о н е ч н а и равна $+\infty$. Аналогичное утверждение справедливо и для нижней грани.

3.5*. Арифметические свойства верхних и нижних граней

Рассмотрим четыре свойства верхних и нижних граней множеств, связанные с арифметическими операциями над числовыми множествами.

Прежде всего определим такие операции.

Арифметической суммой $X_1 + \dots + X_n$ числовых множеств $X_1, \dots, X_n$ называется множество всех чисел x, представимых в виде $x = x_1 + \dots + x_n$, $x_1 \in X_1, \dots, x_n \in X_n$.

Арифметической разностью $X - Y$ числовых множеств X и Y называется множество всех чисел z, представимых в виде $z = x - y$, $x \in X$, $y \in Y$.

Следует, конечно, отличать понятие арифметической суммы $X_1 + \dots + X_n$ и разности $X - Y$ от понятия теоретико-множественных суммы (объединения) $X_1 \cup \dots \cup X_n$ и разности $X \setminus Y$ тех же множеств.

Произведением λX *числа* λ *на числовое множество* X называется множество всех чисел вида λx, $x \in X$.

Произведением XY *двух числовых множеств* X *и* Y называется множество чисел z, представимых в виде $z = xy$, $x \in X$, $y \in Y$.

Первое свойство.

$$\sup(X_1 + \dots + X_n) = \sup X_1 + \dots + \sup X_n, \tag{3.6}$$

$$\inf(X_1 + \dots + X_n) = \inf X_1 + \dots + \inf X_n. \tag{3.7}$$

Второе свойство.

$$\sup(X - Y) = \sup X - \inf Y. \tag{3.8}$$

Докажем первое свойство. Если $x \in X_1 + \dots + X_n$, т. е. $x = x_1 + \dots + x_n$, $x_1 \in X_1, \dots, x_n \in X_n$, то $x_k \leqslant \sup X_k$, $k = 1, 2, \dots, n$, следовательно,

$$x = x_1 + \dots + x_n \leqslant \sup X_1 + \dots + \sup X_n. \tag{3.9}$$

Если

$$y < \sup X_1 + \dots + \sup X_n, \tag{3.10}$$

то для любого $k = 1, 2, \dots, n$ найдется такое число y_k, что

$$y_k < \sup X_k,\ k = 1, 2, \dots, n,\ y_1 + \dots + y_n = y. \tag{3.11}$$

В самом деле, если все верхние грани $\sup X_1, \dots, \sup X_n$ конечны, то можно взять

$$y_k = \sup X_k - \underbrace{\sup X_1 + \dots + \sup X_n - y}_{n}.$$

Тогда $y_1 + \dots + y_n = y$. Если же среди верхних граней $\sup X_1, \dots, \sup X_n$ имеется хотя бы одна равная $+\infty$,

например $X_n = +\infty$, то в качестве $y_1, \ldots, y_{n-1}$ можно взять любые числа $y_i < \sup X_i$, $i = 1, 2, \ldots, n - 1$, а за y_n — число

$$y_n = y - (y_1 + \ldots + y_{n-1}).$$

В этом случае $y_n < +\infty = \sup X_n$ и, очевидно, $y_1 + \ldots + y_n = y$, т. е. снова выполняются условия (3.11).

Из неравенств (3.11) следует, что существуют такие $x_k \in \in X_k$, что

$$y_k < x_k \leqslant \sup X_k,\ k = 1, 2, \ldots, n.$$

Полагая $x = x_1 + \ldots + x_n$, получаем

$$x \in X_1 + \ldots + X_n, \tag{3.12}$$
$$x = x_1 + \ldots + x_n > y_1 + \ldots + y_n \geqslant y.$$

Таким образом, выполняются оба условия определения верхней грани (см. формулы (3.9) и (3.12)), т. е. $\sup X_1 + + \ldots + \sup X_n$ действительно является верхней гранью множества $X_1 + \ldots + X_n$. □

Аналогично доказывается формула (3.7).

Докажем второе свойство. Если $z \in X - Y$, т. е. $z = x - y$, $x \in X$, $y \in Y$, то $x \leqslant \sup X$, $y \geqslant \inf Y$ и, следовательно,

$$z = x - y \leqslant \sup X - \inf Y. \tag{3.13}$$

Если

$$z_z < \sup X - \inf Y, \tag{3.14}$$

то найдутся такие числа

$$z_1 < \sup X,\quad z_2 > \inf Y, \tag{3.15}$$

что $z_1 - z_2 > z$. Такие z_1 и z_2 существуют в силу неравенства (3.14) как в случае, когда грани $\sup X$, $\inf Y$ — конечные, так и в случае, когда одна из них или обе бесконечные.

Из неравенств (3.15) следует, что существуют такие числа $x \in X$ и $y \in Y$, что

$$z_1 < x \leqslant \sup X,\quad z_2 > y \geqslant \inf Y. \tag{3.16}$$

В результате $x - y \in X - Y$ и

$$x - y > z_1 - z_2 > z. \tag{3.17}$$

Таким образом, снова выполняются оба условия определения верхней грани (см. (3.13) и (3.17)), т. е. $\sup X - \inf Y$ действительно является верхней гранью множества $X - Y$. □

Т р е т ь е с в о й с т в о. *Если* $\lambda \geqslant 0$, *то*

$$\sup \lambda X = \lambda \sup X,\quad \inf \lambda X = \lambda \inf X, \tag{3.18}$$

а если $\lambda < 0$, *то*

$$\sup \lambda X = \lambda \inf X, \quad \inf \lambda X = \lambda \sup X. \tag{3.19}$$

Докажем первое из равенств (3.18). Пусть $\lambda > 0$. Если $y \in \lambda X$, т. е. $y \in \lambda x$, где $x \in X$ и, следовательно, $x \leqslant \sup X$, то $y = \lambda x \leqslant \lambda \sup X$. Если $y < \lambda \sup X$, т. е. $\frac{y}{\lambda} < \sup X$, то найдется такое $x \in X$, что $x > \frac{y}{\lambda}$ и, следовательно, $\lambda x > y$, где $\lambda x \in \lambda X$. Таким образом, $\lambda \sup X$ является верхней гранью множества λX, т. е. первое из равенств (3.18) доказано. Аналогично доказывается и второе равенство (3.18).

Пусть теперь $\lambda < 0$. Если $y \in \lambda X$, т. е. $y = \lambda x$, где $x \in X$ и, следовательно, $x \geqslant \inf X$, то $\lambda x \leqslant \lambda \inf X$. Если $y < \lambda \inf X$, т. е. $\frac{y}{\lambda} > \inf X$, то найдется такое $x \in X$, что $x < \frac{y}{\lambda}$, поэтому $\lambda x > y$, где $\lambda x \in \lambda X$. Это означает, что $\lambda \inf X$ является верхней гранью множества λX. Первое равенство (3.19) доказано. Аналогично доказывается и второе равенство (3.19). □

Ч е т в е р т о е с в о й с т в о. *Если все числа, входящие в множества* X *и* Y, *неотрицательны, то*

$$\sup XY = \sup X \sup Y, \quad \inf XY = \inf X \inf Y. \tag{3.20}$$

Доказательство этого свойства проводится тем же методом, что и доказательство предыдущих трех свойств верхних и нижних граней множеств, и предоставляется читателю.

Покажем теперь, что из теоремы о существовании верхних и нижних граней вытекают два важных свойства действительных чисел, так называемые принцип Архимеда[1] и принцип вложенных отрезков.

3.6. Принцип Архимеда

Принцип Архимеда действительных чисел состоит в следующем.

---

[1] Архимед (287—212 до н. э.) — древнегреческий математик и механик.

ТЕОРЕМА 2. *Каково бы ни было действительное число a, существует такое натуральное число n, что $n > a$, т. е.*

na
0 a b
Рис. 8

$$\forall a \in \boldsymbol{R}\ \exists n \in \boldsymbol{N} : n > a. \qquad (3.21)$$

Доказательство. Допустим, что принцип Архимеда не выполняется. Это означает, что существует такое число a, что для всех натуральных n выполняется неравенство $n \leqslant a$, т. е. $\exists a \in \boldsymbol{R}\ \forall n \in \boldsymbol{N} : n \leqslant a$. Это означает, что число a ограничивает сверху множество натуральных чисел. Поэтому множество натуральных чисел, как всякое непустое ограниченное сверху числовое множество, согласно теореме 1 п. 3.4 имеет конечную верхнюю грань. Обозначим ее через β, $\beta = \sup \boldsymbol{N}$.

Так как $\beta - 1 < \beta$, то, согласно свойству 2^0 верхней грани в определении 4 п. 3.4, существует такое натуральное число n, что $n > \beta - 1$. Но тогда $n + 1 > \beta$, причем, согласно определению натуральных чисел, $n + 1 \in \boldsymbol{N}$. Неравенство $n + 1 > \beta$ противоречит тому, что $\beta = \sup \boldsymbol{N}$, так как верхняя грань множества ограничивает его сверху (см. свойство 1^0 верхней грани в определении 4′ п. 3.4). Полученное противоречие показывает, что указанного числа a не существует, т. е. принцип Архимеда справедлив. □

СЛЕДСТВИЕ. *Каковы бы ни были числа a и b, $0 < a < b$, существует такое натуральное число n, что*

$$na > b. \qquad (3.22)$$

Действительно, согласно принципу Архимеда, для числа $\frac{b}{a}$ существует такое натуральное n, что $n > b/a$. Это число искомое, так как, умножая неравенство $n > b/a$ на положительное число a, получаем $na > b$.

Это утверждение имеет простой геометрический смысл: если взять два отрезка соответственно длины a и b, $0 < a < b$, то, последовательно откладывая на большем отрезке от одного из его концов меньший отрезок, мы через конечное число шагов выйдем за пределы большего отрезка (рис. 8).

Пример. Пусть множество X состоит из чисел вида $1/n$, $n = 1, 2, \ldots$. Найдем $\sup X$ и $\inf X$.

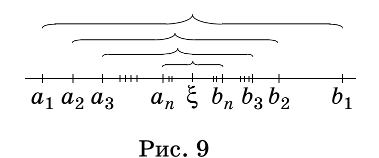


Рис. 9

Множество X имеет наибольшее число 1, поэтому оно и является его верхней гранью: $\sup\limits_{n \in \boldsymbol{N}} \{1/n\} = 1$. Для отыскания нижней грани множества X заметим, что для любого $n = 1, 2, \ldots$ справедливо неравенство $1/n > 0$, т. е. число нуль ограничивает снизу множество X. Покажем, что оно наибольшее среди всех таких чисел. Пусть $\varepsilon > 0$; тогда, согласно принципу Архимеда, существует такое натуральное n, что $n > 1/\varepsilon$, или, что то же самое, $1/n < \varepsilon$. Это неравенство показывает, что любое число $\varepsilon > 0$ уже не ограничивает снизу множество X, ибо $1/n \in X$ при любом $n = 1, 2, \ldots$. Итак, нуль — наибольшее из всех чисел, ограничивающих снизу множество X, т. е. $\inf\limits_{n \in \boldsymbol{N}} \{1/n\} = 0$.

3.7. Принцип вложенных отрезков

Прежде всего поясним, какая система отрезков называется вложенной.

Определение 6. *Система числовых отрезков*

$$[a_1, b_1], [a_2, b_2], \ldots, [a_n, b_n], \ldots\ a_n \in \boldsymbol{R},\ b_n \in \boldsymbol{R},\ n = 1, 2, \ldots,$$

называется системой вложенных отрезков, если

$$a_1 \leqslant a_2 \leqslant \ldots \leqslant a_n \leqslant \ldots \leqslant b_n \leqslant \ldots \leqslant b_2 \leqslant b_1, \tag{3.23}$$

т. е. если каждый следующий отрезок $[a_{n+1}, b_{n+1}]$ *содержится в предыдущем* $[a_n, b_n]$ (рис. 9):

$$[a_1, b_1] \supset [a_2, b_2] \supset \ldots \supset [a_n, b_n] \supset \ldots .$$

ТЕОРЕМА 3. *Для всякой системы вложенных отрезков существует хотя бы одно число, которое принадлежит всем отрезкам данной системы.*

Это свойство действительных чисел называют также *непрерывностью множества действительных чисел в смысле Кантора*[1].

Доказательство. Пусть задана система вложенных отрезков $[a_n, b_n]$, $n = 1, 2 \ldots$. Обозначим через A множество всех левых концов a_n — отрезков этой системы, а через B —

[1] Г. Кантор (1845—1918) — немецкий математик.

множество их правых концов b_n. Для любых номеров m и n выполняется неравенство

$$a_m \leqslant b_n. \qquad (3.24)$$

В самом деле, если $n \geqslant m$, то из неравенств (3.23) следует, что $a_m \leqslant a_n \leqslant b_n$, а если $n < m$, то $a_m \leqslant b_m \leqslant b_n$.

Поэтому из неравенств (3.24), в силу свойства непрерывности действительных чисел, следует, что существует такое число ξ, для которого при всех номерах m и n выполняется неравенство $a_m \leqslant \xi \leqslant b_n$, в частности неравенство $a_n \leqslant \xi \leqslant b_n$, $n = 1, 2, \ldots$. Это и означает, что число ξ принадлежит всем отрезкам $[a_n, b_n]$. □

Приведем условие, при котором пересечение системы вложенных отрезков состоит только из одной точки.

Определение 7. *Пусть задана система отрезков* $[a_n, b_n]$, $a_n \in \boldsymbol{R}$, $b_n \in \boldsymbol{R}$, $n = 1, 2, \ldots$. *Будем говорить, что длина* $b_n - a_n$ *отрезков этой системы стремится к нулю, если для каждого числа* $\varepsilon > 0$ *существует такой номер* n_ε, *что для всех номеров* $n > n_\varepsilon$ *выполняется неравенство*

$$b_n - a_n < \varepsilon. \qquad (3.25)$$

В курсе элементарной математики вводится понятие предела последовательности. Сформулированное определение, если использовать термин предела, означает, что $\lim\limits_{n \to \infty} (b_n - a_n) = 0$. В данном курсе пределу последовательности будет посвящен следующий параграф.

Отметим, что термин «номер» является синонимом термина «натуральное число». Индекс ε у числа n_ε показывает, что это число зависит от задаваемого числа $\varepsilon > 0$.

ТЕОРЕМА 4. *Для всякой системы* $[a_n, b_n]$, $n = 1, 2, \ldots$, *вложенных отрезков с длинами, стремящимися к нулю, существует единственная точка* ξ, *принадлежащая всем отрезкам данной системы* (см. рис. 9), *причем*

$$\xi = \sup_{n \in N} \{a_n\} = \inf_{n \in N} \{b_n\}. \qquad (3.26)$$

Доказательство. Если точки ξ и η принадлежат всем отрезкам рассматриваемой системы, т. е.

$$\xi \in [a_n, b_n], \quad \eta \in [a_n, b_n], \quad n = 1, 2, \ldots ,$$

то для всех номеров n выполняются неравенства

$$|\eta - \xi| \leqslant b_n - a_n,$$

следовательно, в силу условий (3.25), для любого $\varepsilon > 0$ справедливо неравенство

$$|\eta - \xi| < \varepsilon. \tag{3.27}$$

Так как ε — произвольное положительное число, то неравенство (3.27) может иметь место только тогда, когда $\xi = \eta$ (если бы $\xi \neq \eta$, то, например, при $\varepsilon = \frac{1}{2}|\eta - \xi|$ неравенство (3.27) было бы противоречиво). Это означает, что существует единственное число ξ, принадлежащее всем отрезкам $[a_n, b_n]$:

$$a_n \leqslant \xi \leqslant b_n, \quad n = 1, 2, \ldots . \tag{3.28}$$

Из этих неравенств видно, что число ξ ограничивает сверху числа a_n и снизу числа b_n, поэтому, в силу определения верхней и нижней граней, справедливы неравенства

$$a_n \leqslant \sup\{a_n\} \leqslant \xi \leqslant \inf\{b_n\} \leqslant b_n, \quad n = 1, 2, \ldots . \tag{3.29}$$

Если бы, например, оказалось, что $\sup\{a_n\} < \xi$, то любое число η такое, что $\sup\{a_n\} < \eta < \xi$, в силу неравенства (3.29), для всех $n = 1, 2, \ldots$ удовлетворяло бы условию $a_n \leqslant \eta \leqslant b_n$, т. е. также бы принадлежало всем отрезкам $[a_n, b_n]$, что невозможно, так как $\eta \neq \xi$. Аналогично доказывается невозможность неравенства $\xi < \inf\{b_n\}$. Итак, соотношение (3.26) доказано. □

Очень часто в различных доказательствах применяется следующая конструкция построения системы вложенных отрезков с длинами, стремящимися к нулю. Берется отрезок $[a, b]$ и точкой $\frac{a+b}{2}$ делится на два равных отрезка $\left[a; \frac{a+b}{2}\right]$ и $\left[\frac{a+b}{2}; b\right]$ длины $\frac{b-a}{2}$. Далее выбирается один из этих отрезков (какой именно — это зависит от условий конкретной задачи), обозначается через $[a_1, b_1]$ и снова своей серединой делится на два равных отрезка, один из которых обозначается $[a_2, b_2]$ и т. д. В результате получается система вложенных отрезков $[a_n, b_n]$, $n = 1, 2, \ldots$, с длинами $b_n - a_n = \frac{b-a}{2^n}$.

Покажем, что эти длины стремятся к нулю.

Действительно, для всякого $\varepsilon > 0$, согласно принципу Архимеда, найдется такое натуральное n_ε, что $n_\varepsilon > \frac{b-a}{\varepsilon}$, но тогда и для всех $n \geqslant n_\varepsilon$ будет выполняться неравенство $n > \frac{b-a}{\varepsilon}$ и, следовательно, неравенство $\frac{b-a}{n} < \varepsilon$. Замечая, что

$$2^n = (1+1)^n = 1 + n + \frac{n(n-1)}{2} + \ldots > n, \tag{3.30}$$

получаем $\frac{1}{2^n} < \frac{1}{n}$ для $n = 1, 2, \ldots$. Поэтому для всех $n > n_\varepsilon$ справедливо неравенство $\frac{b-a}{2^n} < \frac{b-a}{n} < \varepsilon$. Это и означает стремление к нулю длин отрезков $[a_n, b_n]$ при возрастании n.

Заметим, что принцип вложенных отрезков является свойством, присущим именно множеству действительных чисел. Так, поле одних только рациональных чисел уже не обладает аналогичным свойством.

Например, если взять последовательности «рациональных отрезков» [1; 2], [1,4; 1,5], [1,41; 1,42], [1,414; 1,415][1] ... , т. е. последовательность множеств рациональных чисел, лежащих на отрезках, концы a_n и b_n, $n = 1, 2, \ldots$ которых являются значениями $\sqrt{2}$, вычисленными соответственно с недостатком и с избытком с точностью $1/10^n$, $n = 0, 1, 2, \ldots$[2], то, очевидно, не существует никакого рационального числа, принадлежащего всем этим отрезкам. В самом деле, таким числом могло быть только число $\sqrt{2}$ (почему?), которое, однако, не является рациональным[3].

Можно доказать и более точное утверждение. Назовем поле *архимедовым*, если для него выполняется принцип Архимеда, т. е. справедливо утверждение теоремы 2 из п. 3.6. Свойство упорядоченного поля (определение поля см. в замечании в конце п. 2.2*), состоящее в том, что для его элемен-

[1] В том случае, когда концы отрезка $[a, b]$ записаны в виде десятичной дроби, запятая между a и b заменяется точкой с запятой.

[2] Это означает, что $a_n^2 \leqslant 2 \leqslant b_n^2$ и $b_n - a_n = \frac{1}{10^n}$, $n = 0, 1, 2, \ldots$.

[3] Доказательство иррациональности числа $\sqrt{2}$, обычно проводимое в элементарной математике, воспроизведено ниже в п. 6.3.

тов выполняется свойство V из п. 2.1, называется *непрерывностью поля по Дедекинду* (см. также п. 2.5*), а свойство упорядоченного поля, выражающееся в том, что каждая система его вложенных отрезков имеет непустое пересечение, — *непрерывностью поля по Кантору*.

Для архимедовых упорядоченных полей можно показать, что их непрерывность по Дедекинду, непрерывность по Кантору и существование конечной верхней грани у каждого непустого ограниченного сверху множества эквивалентны между собой, т. е. из любого из этих свойств, принятого за аксиому, вытекают остальные два.

Было показано, что из непрерывности по Дедекинду следует существование конечной верхней грани у ограниченного сверху множества, откуда, в свою очередь, следует непрерывность по Кантору. Для того чтобы завершить доказательство указанной эквивалентности трех понятий непрерывности архимедовых полей, достаточно показать, что из непрерывности по Кантору следует непрерывность по Дедекинду. Доказательство этого утверждения можно найти, например, в книге: *Кудрявцев Л. Д.* Математический анализ. Т. 1. М., 1973.

Отметим еще следующее: при доказательстве утверждения, что упорядоченное поле, непрерывное по Дедекинду, является непрерывным и по Кантору, требование архимедовости поля можно отбросить. Действительно, в п. 3.6 было доказано, что из непрерывности упорядоченного поля по Дедекинду, т. е. из наличия свойства V, сформулированного в п. 2.1, следует выполнение принципа Архимеда.

В дальнейшем упорядоченное по Дедекинду поле будем просто для краткости называть *непрерывным*.

В заключение обратим внимание на то, что утверждение, аналогичное теореме 3, оказывается уже неверным для числовых промежутков других типов, чем отрезки. Например, система вложенных интервалов $\left(0, \frac{1}{n}\right)$, $n = 1, 2, \ldots$: каждый последующий интервал содержится в предыдущем, т. е.

$$\left(0, \frac{1}{n+1}\right) \subset \left(0, \frac{1}{n}\right), \; n = 1, 2, \ldots,$$

имеет пустое пересечение. Но, конечно, существуют и такие системы вложенных интервалов, которые имеют непустое пересечение.

3.8*. Единственность непрерывного упорядоченного поля

Множество действительных чисел является в некотором смысле единственным непрерывным упорядоченным полем, точнее единственным с точностью до изоморфизма. Разъясним, что это означает.

Два поля $\mathcal{P}$ и $\mathcal{P}^*$ называются *изоморфными*, если существует такое взаимно однозначное отображение f поля $\mathcal{P}$ на поле $\mathcal{P}^*$, что для любых двух элементов $x \in \mathcal{P}$ и $y \in \mathcal{P}$ выполняются условия

$$f(x + y) = f(x) + f(y), \ f(xy) = f(x)f(y).$$

Отображение f называется в этом случае *изоморфизмом* или изоморфным отображением.

Короче, два поля называются изоморфными, если существует взаимно однозначное отображение одного из них на другое (биекция), сохраняющее сложение и умножение их элементов.

Если поля $\mathcal{P}$ и $\mathcal{P}^*$ упорядоченные и существует изоморфное отображение f поля $\mathcal{P}$ на поле $\mathcal{P}^*$, сохраняющее отношение порядка, т. е. такое, что для любых $x \in \mathcal{P}$ и $y \in \mathcal{P}$, для которых $x < y$, имеет место соотношение $f(x) < f(y)$, то поля $\mathcal{P}$ и $\mathcal{P}^*$ называются *изоморфными упорядоченными полями*.

Докажем, что множество действительных чисел однозначно определяется системой аксиом I—V (см. п. 2.1) с точностью до природы элементов.

ТЕОРЕМА 5. *Все непрерывные упорядоченные поля изоморфны между собой.*

Доказательство. Пусть даны два непрерывных упорядоченных поля $\boldsymbol{R}$ и $\boldsymbol{R}^*$. Будем последовательно устанавливать соответствие между их элементами $x \in \boldsymbol{R}$ и $x^* \in \boldsymbol{R}^*$. Это соответствие обозначим через f. В каждом из полей $\boldsymbol{R}$ и $\boldsymbol{R}^*$ есть нуль и единица, соответственно 0 и 1 в $\boldsymbol{R}$ и 0^* и 1^* в $\boldsymbol{R}^*$. Поставим в соответствие нулю 0 нуль 0^*, а единице 1 — единицу 1^*:

$$f(0) \overset{\text{def}}{=} 0^*, \ f(1) \overset{\text{def}}{=} 1^*. \qquad (3.31)$$

Элементу $n = \underbrace{1 + 1 + \ldots + 1}_{n \text{ раз}}$ поставим в соответствие элемент $\underbrace{f(1) + f(1) + \ldots + f(1)}_{n \text{ раз}} = \underbrace{1^* + 1^* + \ldots + 1^*}_{n \text{ раз}}$, таким образом,

$$f(n) \overset{\text{def}}{=} \underbrace{1^* + 1^* + \ldots + 1^*}_{n \text{ раз}}, \tag{3.32}$$

а элементу $-n$ элемент $-f(n)$

$$f(-n) = -f(n). \tag{3.33}$$

Установленное соответствие f является взаимно однозначным, при этом каждый целый элемент поля $\boldsymbol{R}^*$, т. е. его нуль 0^*, элемент вида $n^* = 1^* + 1^* + \ldots + 1^*$ или $-n^*$, оказывается, в силу равенств (3.31)—(3.33), поставленным в соответствие какому-то целому элементу поля $\boldsymbol{R}$. Таким образом, отображение f устанавливает взаимно однозначное соответствие между множествами $\boldsymbol{Z}$ и $\boldsymbol{Z}^*$ целых элементов полей $\boldsymbol{R}$ и $\boldsymbol{R}^*$. Это соответствие сохраняет отношение порядка: если $0 < m < n$, то

$$f(m) = \underbrace{1^* + 1^* + \ldots + 1^*}_{m \text{ раз}} < \underbrace{1^* + 1^* + \ldots + 1^*}_{n \text{ раз}} = f(n). \tag{3.34}$$

В силу соответствия (3.33), отношение порядка сохраняется и для всех целых элементов любого знака.

Сохраняются и операции сложения и умножения: для любых $m \in \boldsymbol{Z}$ и $n \in \boldsymbol{Z}$ имеют место следующие равенства:

$$f(m + n) = f(m) + f(n), \tag{3.35}$$

$$f(mn) = f(m)f(n). \tag{3.36}$$

Действительно, если $m > 0$ и $n > 0$, то

$$f(m + n) = \underbrace{1^* + \ldots + 1^*}_{m + n \text{ раз}} = \underbrace{1^* + \ldots + 1^*}_{m \text{ раз}} + \underbrace{1^* + \ldots + 1^*}_{n \text{ раз}} = f(m) + f(n). \tag{3.37}$$

Если же $m < 0$ и $0 < -m < n$, то

$$f(m + n) = \underbrace{1^* + \ldots + 1^*}_{n + m = n - (-m) \text{ раз}} = -\underbrace{(1^* + \ldots + 1^*)}_{-m \text{ раз}} + \underbrace{(1^* + \ldots + 1^*)}_{n \text{ раз}} = -f(-m) + f(n) \underset{(3.33)}{=} f(m) + f(n). \tag{3.38}$$

Случай $0 < m < -n$ сводится к предыдущему:

$$f(m+n) \underset{(3.33)}{=} -f(-m-n) \underset{(3.38)}{=} -(f(-m)+$$
$$+f(-n)) \underset{(3.33)}{=} f(m)+f(n).$$

Наконец, если $m < 0$, $n < 0$, то

$$f(m+n) \underset{(3.33)}{=} -f(-m-n) \underset{(3.37)}{=} -(f(-m)+f(-n)) \underset{(3.33)}{=} f(m)+f(n).$$

Если $m = 0$ или $n = 0$, то соотношения (3.35) и (3.36) очевидны. Для доказательства соотношения (3.36) заметим, что в полях $\boldsymbol{R}$ и $\boldsymbol{R}^*$, как и вообще во всех упорядоченных полях, имеется понятие абсолютной величины элементов и для их умножения справедливы соотношения:

$$\text{если } x \geqslant 0,\ y \geqslant 0 \text{ или } x \leqslant 0,\ y \leqslant 0, \text{ то } xy = |xy|, \tag{3.39}$$
$$\text{если } x > 0,\ y < 0 \text{ или } x < 0,\ y > 0, \text{ то } xy = -|xy|. \tag{3.40}$$

Из равенств (3.32) и (3.33) следует, что для каждого $n \in \boldsymbol{Z}$ имеет место равенство

$$f(|n|) = |f(n)|. \tag{3.41}$$

Если для сомножителей m, $n \in \boldsymbol{Z}$ имеет место случай (3.39), то, в силу (3.33), такой же случай имеет место и для сомножителей $f(m)$, $f(n)$, поэтому

$$f(mn) \underset{(3.39)}{=} f(|m||n|) \underset{(3.32)}{=} \underbrace{1^* + \ldots + 1^*}_{|m|\,|n| \text{ раз}} =$$
$$= \underbrace{(1^* + \ldots + 1^*)}_{|m| \text{ раз}} \underbrace{(1^* + \ldots + 1^*)}_{|n| \text{ раз}} \underset{(3.32)}{=}$$
$$= f(|m|)f(|n|) \underset{(3.41)}{=} |f(m)f(n)| \underset{(3.39)}{=} f(m)f(n). \tag{3.42}$$

Если же сомножители m и n разного знака, то, в силу (3.32) и (3.33), $f(m)$ и $f(n)$ также имеют разные знаки; следовательно,

$$f(mn) \underset{(3.40)}{=} f(-|m||n|) \underset{(3.33)}{=} -f(|m||n|) \underset{(3.42)}{=}$$
$$= -f(|m|)f(|n|) \underset{(3.41)}{=} -|f(m)f(n)| \underset{(3.40)}{=} f(m)f(n).$$

Итак, отображение f взаимно однозначно отображает множество $\boldsymbol{Z}$ на $\boldsymbol{Z}^*$, сохраняет отношение порядка и операции сложения и умножения.

Поставим далее каждому рациональному элементу поля $\boldsymbol{R}$, т. е. элементу вида $\frac{m}{n}$, где $m \in \boldsymbol{Z}$, $n \in \boldsymbol{Z}$, $n \neq 0$, элемент $\frac{f(m)}{f(n)}$ поля $\boldsymbol{R}^*$, т. е. положим

$$f\left(\frac{m}{n}\right) \stackrel{\text{def}}{=} \frac{f(m)}{f(n)}. \tag{3.43}$$

Теперь (как это легко проверить) f отображает взаимно однозначно поле $\boldsymbol{Q}$ всех рациональных элементов поля $\boldsymbol{R}$ на поле $\boldsymbol{Q}^*$ всех рациональных элементов поля $\boldsymbol{R}^*$. Это соответствие является изоморфизмом упорядоченных полей $\boldsymbol{Q}$ и $\boldsymbol{Q}^*$. В самом деле, если

$$0 < \frac{m}{n} < \frac{p}{q},\ m,\ n,\ p,\ q \in \boldsymbol{Z},\ n > 0,\ q > 0, \tag{3.44}$$

то $mq < np$, а тогда, в силу (3.34), $f(mq) < f(np)$, откуда $f(m)f(q) \underset{(3.36)}{<} f(n)f(p)$ и, следовательно, $\frac{f(m)}{f(n)} < \frac{f(p)}{f(q)}$, т. е., согласно (3.43),

$$f\left(\frac{m}{n}\right) < f\left(\frac{p}{q}\right). \tag{3.45}$$

Для рациональных элементов с произвольными знаками сохранение отношения порядка при отображении f следует из (3.44)—(3.45) и того, что

$$f\left(-\frac{m}{n}\right) = f\left(\frac{-m}{n}\right) \underset{(3.43)}{=} \frac{f(-m)}{f(n)} \underset{(3.33)}{=} \frac{-f(m)}{f(n)} = -f\left(\frac{m}{n}\right). \tag{3.46}$$

Далее, для любых $\frac{m}{n} \in \boldsymbol{Q}$ и $\frac{p}{q} \in \boldsymbol{Q}$ имеем

$$f\left(\frac{m}{n} + \frac{p}{q}\right) = f\left(\frac{mq + np}{nq}\right) \underset{(3.43)}{=} \frac{f(mq + np)}{f(nq)} \underset{\substack{(3.35)\\(3.36)}}{=} \frac{f(m)f(q) + f(n)f(p)}{f(n)f(q)} =$$

$$= \frac{f(m)}{f(n)} + \frac{f(p)}{f(q)} \underset{(3.43)}{=} f\left(\frac{m}{n}\right) + f\left(\frac{p}{q}\right) \tag{3.47}$$

и, наконец,

$$f\left(\frac{m}{n}\,\frac{p}{q}\right) = f\left(\frac{mp}{nq}\right) \underset{(3.43)}{=} \frac{f(mp)}{f(nq)} \underset{(3.36)}{=} \frac{f(m)f(p)}{f(n)f(q)} =$$

$$= \frac{f(m)}{f(n)}\,\frac{f(p)}{f(q)} \underset{(3.43)}{=} f\left(\frac{m}{n}\right)f\left(\frac{p}{q}\right). \tag{3.48}$$

Иррациональные элементы, т. е. элементы, не являющиеся рациональными, определяются сечениями в областях рациональных элементов, причем, в силу изоморфизма меж-

ду множествами $\boldsymbol{Q}$ и $\boldsymbol{Q}^*$ рациональных элементов, между их сечениями также существует взаимно однозначное соответствие: если $A|B$ — сечение в $\boldsymbol{Q}$, то $f(A)|f(B)$ — сечение в $\boldsymbol{Q}^*$ и мы положим

$$f(A|B) = f(A)|f(B). \tag{3.49}$$

Теперь взаимно однозначное соответствие установлено между всеми элементами полей $\boldsymbol{R}$ и $\boldsymbol{R}^*$. Покажем, что оно также сохраняет отношение порядка и операции сложения и умножения элементов, т. е. является изоморфизмом полей $\boldsymbol{R}$ и $\boldsymbol{R}^*$. Для этого заметим, что (см. п. 3.4)

$$A|B = \sup A = \inf B,\ f(A)|f(B) = \sup f(A) = \inf f(B). \tag{3.50}$$

Если для заданного множества $X \subset \boldsymbol{Q}$ обозначить через A множество таких рациональных чисел r, что для каждого $r \in A$ существует $x \in X$, для которого $r \leqslant x$, и положить $B = \boldsymbol{Q} \backslash A$, то множества A и B образуют сечение $A|B$ в поле $\boldsymbol{Q}$ и

$$\sup X = \sup A = A|B.$$

В силу же изоморфизма полей $\boldsymbol{Q}$ и $\boldsymbol{Q}^*$ имеем

$$\sup f(X) = \sup f(A) = f(A)|f(B) \underset{(3.49)}{=} f(A|B) = f(\sup X).$$

Подобным же образом доказывается аналогичное соотношение для нижних граней. Следовательно, имеют место равенства

$$\sup f(X) = f(\sup X),\ \inf f(X) = f(\inf X),\ X \subset \boldsymbol{Q}. \tag{3.51}$$

Пусть теперь $A|B$ и $C|D$ — сечения в поле $\boldsymbol{Q}$ и

$$A|B \leqslant C|D, \tag{3.52}$$

тогда $A \subset C$ и, следовательно, $f(A) \subset f(C)$, откуда вытекает, что

$$f(A)|f(B) \leqslant f(C)|f(D). \tag{3.53}$$

Далее, для любых сечений $A|B$ и $C|D$ поля $\boldsymbol{Q}$ имеем

$$f(A|B + C|D) \underset{(3.50)}{=} f(\sup A + \sup C) \underset{(3.6)}{=} f(\sup (A + C)) \underset{(3.51)}{=}$$
$$\underset{(3.51)}{=} \sup f(A + C) \underset{(3.47)}{=} \sup (f(A) + f(C)) \underset{(3.6)}{=}$$
$$\underset{(3.6)}{=} \sup f(A) + \sup f(C) \underset{(3.50)}{=} f(A)|f(B) + f(C)|f(D).$$

Аналогично для $A|B \geqslant 0$, $C|D \geqslant 0$ имеем

$$f(A|B \cdot C|D) \underset{(3.50)}{=} f(\inf B \cdot \inf D) = f(\inf BD) \underset{(3.51)}{=}$$

$$\underset{(3.51)}{=} \inf f(BD) \underset{(3.36)}{=} \inf (f(B)f(D)) =$$

$$= \inf f(B) \inf f(D) \underset{(3.50)}{=} f(A)|f(B) \cdot f(C)|f(D). \qquad (3.54)$$

Соотношение

$$f(A|B \cdot C|D) = f(A|B)f(C|D)$$

для произвольных сечений следует из формул (3.39), (3.40), (3.54) и равенства

$$f(-(A|B)) \underset{(3.50)}{=} f(-\sup A) \underset{(3.5)}{=} f(\inf (-A)) \underset{(3.51)}{=} \inf f(-A) \underset{(3.46)}{=}$$

$$\underset{(3.46)}{=} \inf (-f(A)) \underset{(3.5)}{=} -\sup f(A) \underset{(3.50)}{=} -(f(A)|f(B)).$$

Изоморфизм $\boldsymbol{R}$ и $\boldsymbol{R}^*$ доказан. □

В заключение докажем, что множество действительных чисел нельзя расширить до большего непрерывного упорядоченного поля с сохранением соотношения упорядоченности и операций сложения и умножения. Предварительно сформулируем одно определение.

Подмножество поля $\mathcal{P}$ называется *подполем поля* $\mathcal{P}$, если это подмножество само является полем в силу операции сложения и умножения элементов поля $\mathcal{P}$. Подполе поля называется собственным, если оно не совпадает со всем полем.

Подполе упорядоченного поля с соотношением упорядоченности, порожденным соотношением упорядоченности поля, называется *упорядоченным подполем*.

Например, поле всех рациональных чисел является собственным упорядоченным подполем поля всех действительных чисел.

Заметим, что нуль и единица подполя некоторого поля являются нулем и единицей поля.

В самом деле, пусть $\mathcal{P}$ — подполе поля $\mathcal{P}^*$, а 0 и 0^* — соответственно нули подполя $\mathcal{P}$ и поля $\mathcal{P}^*$. Обозначим через a элемент, противоположный элементу 0 в поле $\mathcal{P}^*$:

$$0 + a = 0^*. \qquad (3.55)$$

В подполе $\mathcal{P}$ имеет место равенство $0 + 0 = 0$. Прибавим к обеим его частям элемент $a : 0 + (0 + a) = 0 + a$, отсюда, в силу (3.55), получаем $0 + 0^* = 0^*$. Но левая часть этого равенства, согласно определению нуля 0^* поля $\mathcal{P}^*$, равна нулю; следовательно, $0 = 0^*$.

Аналогично, если 1 и 1^* — соответственно единицы подполя $\mathcal{P}$ и поля $\mathcal{P}^*$, то из равенства $1 \cdot 1 = 1$, умножая его на элемент 1^{-1}, обратный в поле $\mathcal{P}^*$ элементу 1: $1 \cdot 1^{-1} = 1^*$, получаем $1 \cdot 1^* = 1^*$, но $1 \cdot 1^* = 1$, следовательно, $1 = 1^*$.

ТЕОРЕМА 6. *Не существует непрерывного упорядоченного поля, содержащего в себе поле действительных чисел в качестве собственного упорядоченного подполя.*

СЛЕДСТВИЕ. *Множество действительных чисел не изоморфно никакой своей собственной части.*

Доказательство. Пусть $\boldsymbol{R}^*$ — непрерывное упорядоченное поле, содержащее в себе поле действительных чисел $\boldsymbol{R}$ как упорядоченное подполе. Покажем, что из этого следует, что $\boldsymbol{R}^* = \boldsymbol{R}$.

При доказательстве теоремы 5 было показано, что между двумя любыми упорядоченными полями, а следовательно, в частности, между полями $\boldsymbol{R}$ и $\boldsymbol{R}^*$, можно установить изоморфизм f, продолжив соответствующим образом отображение

$$f(0) = 0^*,\ f(1) = 1^* \tag{3.56}$$

на все элементы поля $\boldsymbol{R}$, где 0 и 1 — нуль и единица поля $\boldsymbol{R}$, а 0^* и 1^* — нуль и единица поля $\boldsymbol{R}^*$.

Согласно сказанному выше, из того, что поле $\boldsymbol{R}$ является подполем $\boldsymbol{R}^*$, вытекает, что $0 = 0^*$ и $1 = 1^*$, и, следовательно, равенства (3.56) превращаются в равенства

$$f(0) = 0, f(1) = 1. \tag{3.57}$$

Покажем, что из этих соотношений вытекает, что и для любого элемента $x \in \boldsymbol{R}$ имеет место равенство

$$f(x) = x. \tag{3.58}$$

В самом деле, если n — натуральное число, то из формул (3.32) и (3.57) следует, что

$$f(n) = 1 + 1 + 1 + \ldots + 1 = n. \tag{3.59}$$

Если n — целое число, то из формул (3.33) и (3.59) имеем также

$$f(n) = n. \tag{3.60}$$

Далее, если $x = \frac{m}{n}$ — рациональное число, то

$$f\left(\frac{m}{n}\right) \underset{(3.43)}{=} \frac{f(m)}{f(n)} = \frac{m}{n}.$$

Все рациональные числа при отображении f остаются на месте, поэтому для любого сечения $A|B$ в области рациональных чисел имеем

$$f(A) = A, \ f(B) = B \tag{3.61}$$

и, следовательно,

$$f(A|B) \underset{(3.49)}{=} f(A)|f(B) \underset{(3.61)}{=} A|B. \tag{3.62}$$

Каждое действительное число определяет два сечения в области рациональных чисел. Для данного числа $x \in \boldsymbol{R}$ обозначим через $A|B$ одно из таких сечений: $x = A|B$. Тогда будем иметь

$$f(x) = f(A|B) \underset{(3.62)}{=} A|B = x.$$

Итак, равенство (3.58) доказано для всех действительных чисел x.

Отображение f, будучи изоморфизмом полей $\boldsymbol{R}$ и $\boldsymbol{R}^*$, является отображением поля $\boldsymbol{R}$ на все поле $\boldsymbol{R}^*$, поэтому соотношение (3.58) означает, что f является тождественным отображением поля $\boldsymbol{R}$ на поле $\boldsymbol{R}^*$, т. е. $\boldsymbol{R}^* = \boldsymbol{R}$. □

§ 4.

Предел числовой последовательности

4.1. Определение предела числовой последовательности

Одним из важнейших понятий математического анализа является понятие предела. Мы начнем его изучение с предела последовательности действительных чисел (определение последовательности см. в п. 1.3*). Под бесконечно удаленной

точкой числовой прямой будем понимать одну из бесконечностей $+\infty$, $-\infty$ или ∞ (см. п. 3.1).

Определение 1. *Точка a (конечная или бесконечно удаленная) числовой прямой называется пределом некоторой числовой последовательности действительных чисел, если, какова бы ни была окрестность точки a, она содержит все члены рассматриваемой последовательности начиная с некоторого номера.*

Этот номер зависит, вообще говоря, от выбора окрестности точки a.

Сформулированное условие равносильно тому, что вне любой окрестности точки a находится лишь конечное множество членов рассматриваемой последовательности, в частности ни одного (т. е. пустое множество, которое причисляется к конечным множествам).

Вспомнив, что окрестности конечных и бесконечно удаленных точек числовой прямой определяются заданием некоторого числа $\varepsilon > 0$ (см. п. 3.2), определение предела последовательности действительных чисел можно перефразировать следующим образом.

Точка a (конечная или бесконечно удаленная) числовой прямой называется пределом последовательности $\{x_n\}$ действительных чисел, если для любого $\varepsilon > 0$ существует такой номер n_ε, что для всех номеров $n > n_\varepsilon$ члены x_n содержатся в окрестности $U(a; \varepsilon)$:

$$x_n \in U(a; \varepsilon).$$

Если выполняется это условие, то пишут $\lim\limits_{n \to \infty} x_n = a$ или $x_n \to a$ при $n \to \infty$ и говорят, что члены последовательности $\{x_n\}$ стремятся к a.

С помощью логических символов существования и всеобщности определение предела записывается следующим образом:

$$a = \lim_{n \to \infty} x_n \overset{\text{def}}{\Leftrightarrow} \exists n_\varepsilon \, \forall \, n > n_\varepsilon : x_n \in U(a; \varepsilon)$$

(здесь и в дальнейшем буквой n, быть может с тем или иным индексом, всегда будем обозначать натуральное число, если специально не оговорено что-либо другое).

Индекс ε у номера n_ε подчеркивает, что этот номер зависит, вообще говоря, от выбранного $\varepsilon > 0$. Эта зависимость отражена, конечно, уже в самой формулировке определения предела, поэтому ее можно и не отражать в записи (что было сделано лишь для большей наглядности). Действительно, часто вместо n_ε пишут, например, $n_0 \in \boldsymbol{N}$ или $N \in \boldsymbol{N}$.

Если предел последовательности действительных чисел является конечной точкой числовой прямой, т. е. числом, то говорят, что последовательность имеет конечный предел.

Определение 2. *Если числовая последовательность имеет конечный предел, то она называется сходящейся.*

Если использовать логические символы, то это определение можно записать следующим образом:

$$\exists\, a \in \boldsymbol{R}\ \forall\, \varepsilon > 0\ \exists\, n_\varepsilon \in \boldsymbol{N}\ \forall\, n > n_\varepsilon : \{x_n - a| < \varepsilon.$$

Последовательность, не являющаяся сходящейся, называется расходящейся.

Для случая конечного предела определение 1 предела можно перефразировать следующим образом.

Число a является пределом последовательности $\{x_n\}$ действительных чисел, если для любого $\varepsilon > 0$ существует такой номер n_ε, что для всех номеров $n > n_\varepsilon$ выполняется неравенство

$$|x_n - a| < \varepsilon. \qquad (4.1)$$

С помощью логических символов это определение записывается следующим образом:

$$a = \lim_{n \to \infty} x_n \overset{\text{def}}{\Leftrightarrow} \forall\, \varepsilon > 0\ \exists\, n_\varepsilon\ \forall\, n > n_\varepsilon : |x_n - a| < \varepsilon.$$

Очевидно, что неравенство (4.1) равносильно неравенству

$$a - \varepsilon < x_n < a + \varepsilon.$$

Если $\lim\limits_{n \to \infty} x_n = a$ и $x_n < a$ (соответственно $x_n > a$) для всех $n = 1, 2, \ldots$, то говорят, что последовательность $\{x_n\}$ сходится к числу *a слева* (соответственно *справа*), и иногда вместо $\lim\limits_{n \to \infty} x_n = a$ в этом случае пишут $\lim\limits_{n \to \infty} x_n = a - 0$ (соответственно $\lim\limits_{n \to \infty} x_n = a + 0$).

В том случае, когда $a = 0$, вместо $0 + 0$ и $0 - 0$ пишут соответственно просто $+0$ и -0.

Сформулируем на языке ε определение предела числовой последовательности в том случае, когда этот предел является той или иной бесконечно удаленной точкой (или, как говорят, равен бесконечности).

Например, ∞ является пределом последовательности $\{x_n\}$, если для любого $\varepsilon > 0$ существует такой номер n_ε, что для всех номеров $n > n_\varepsilon$ выполняется включение $x_n \in U(\infty; \varepsilon)$, или, что то же самое, неравенство $|x_n| > \frac{1}{\varepsilon}$. С помощью логических символов это утверждение записывается следующим образом:

$$\lim_{n \to \infty} x_n = \infty \overset{\text{def}}{\Leftrightarrow} \forall\, \varepsilon > 0\; \exists\, n_\varepsilon\; \forall\, n > n_\varepsilon : |x_n| > \frac{1}{\varepsilon}.$$

Аналогично определение предела последовательности перефразируется для случая, когда этот предел равен бесконечности с определенным знаком. Для краткости ограничимся записью этих определений только с помощью логических символов:

$$\lim_{n \to \infty} x_n = +\infty \overset{\text{def}}{\Leftrightarrow} \forall\, \varepsilon > 0\; \exists\, n_\varepsilon\; \forall\, n > n_\varepsilon : x_n > \frac{1}{\varepsilon};$$

$$\lim_{n \to \infty} x_n = -\infty \overset{\text{def}}{\Leftrightarrow} \forall\, \varepsilon > 0\; \exists\, n_\varepsilon\; \forall\, n > n_\varepsilon : x_n < -\frac{1}{\varepsilon}.$$

Очевидно, что если $\lim\limits_{n \to \infty} x_n = +\infty$ или $\lim\limits_{n \to \infty} x_n = -\infty$, то и $\lim\limits_{n \to \infty} x_n = \infty$.

Определение 3. *Последовательность, пределом которой является бесконечность, называется бесконечно большой.*

Понятие конечного предела последовательности связано в определенном смысле с встречающейся на практике задачей получения значения некоторой интересующей нас величины с наперед заданной фиксированной точностью $\varepsilon > 0$. Последовательные приближенные значения x_n рассматриваемой величины могут получаться в результате проведения каких-либо экспериментов или вычисления по каким-нибудь рекуррентным формулам или каким-то другим путем. Эта задача будет, очевидно, решена, если найдется номер n_ε, начиная с которого все значения x_n будут отклоняться от точного значения рассматриваемой величины в пределах заданной точности. Конечно, если указанное n_ε существует

лишь для одного данного $\varepsilon > 0$, это еще не означает, что последовательность $\{x_n\}$ сходится: в определении предела последовательности требуется, чтобы соответствующий номер n_ε можно было подобрать для любого $\varepsilon > 0$.

В дальнейшем всегда под пределом последовательности будем понимать конечный предел, т. е. число, если, конечно, не оговорено противное.

Примеры. 1. Последовательность $\left\{\frac{1}{n}\right\}$ сходится и имеет своим пределом нуль. В самом деле, каково бы ни было $\varepsilon > 0$, согласно принципу Архимеда (см. п. 3.6) действительных чисел, существует такое натуральное число n_ε, что $n_\varepsilon > \frac{1}{\varepsilon}$. Поэтому для всех $n > n_\varepsilon$ выполняется неравенство $0 < \frac{1}{n} < \frac{1}{n_\varepsilon} < \varepsilon$, а это и означает, что $\lim\limits_{n \to \infty} \frac{1}{n} = 0$. Последовательность $\left\{\frac{1}{n}\right\}$ сходится к нулю справа.

2. Последовательность $\left\{\sin \frac{\pi}{2} n\right\}$ является расходящейся. В самом деле, каково бы ни было число a, вне его ε-окрестности, например при $0 < \varepsilon < 1$, заведомо лежит бесконечное число членов данной последовательности, и, значит, оно не является ее пределом.

3. Последовательность $\left\{\frac{1}{n} \sin \frac{\pi}{2} n\right\}$ сходится и

$$\lim_{n \to \infty} \frac{1}{n} \sin \frac{\pi}{2} n = 0.$$

Это следует (почему?) из того, что

$$\left|\frac{1}{n} \sin \frac{\pi}{2} n\right| \leqslant \frac{1}{n} \text{ и } \lim_{n \to \infty} \frac{1}{n} = 0.$$

Сходящаяся последовательность $\left\{\frac{1}{n} \sin \frac{\pi}{2} n\right\}$ не является последовательностью, сходящейся к своему пределу слева или справа.

4. Последовательность $\{n\}$ расходится. Действительно, каково бы ни было число a, для любого $\varepsilon > 0$, в частности для $\varepsilon = 1$, найдется, согласно принципу Архимеда, такое натуральное n_0, что $n_0 > a + 1$. Следовательно, и для всех нату-

ральных $n > n_0$ имеем $n > a + 1$. Поэтому никакое число a не может являться пределом последовательности $\{n\}$.

В примерах 2 и 4 при доказательстве расходимости последовательностей было использовано *позитивное определение* того обстоятельства, что некоторое число a не являлось пределом данной последовательности.

Определение 4. *Конечная или бесконечно удаленная точка a числовой прямой* $\boldsymbol{R}$ *не является*[1] *пределом последовательности* $x_n \in \boldsymbol{R}$, $n = 1, 2, \dots$, *если существует такое* $\varepsilon > 0$, *что для всякого натурального n существует такое натуральное* $m > n$, *что* $x_m \notin U(a; \varepsilon)$.

С помощью логических символов это определение записывается следующим образом:

$$\lim_{n \to \infty} x_n \neq a \Leftrightarrow \exists\, \varepsilon > 0\ \forall\, n \in \boldsymbol{N}\ \exists\, m > n : x_m \notin U(a; \varepsilon).$$

Напомним, что при формулировании отрицания какого-либо утверждения логические символы существования $\exists$ и всеобщности $\forall$ меняются местами. Именно так и произошло в данном случае, в чем легко убедиться, сравнив запись определений 1 и 4, записанную с помощью логических символов.

Заметим, что определение 4 не является самостоятельным определением — оно является логическим следствием определения 1.

УПРАЖНЕНИЕ 1. Сформулировать позитивное определение понятия расходящейся последовательности.

УПРАЖНЕНИЕ 2. Доказать, что если $\lim\limits_{n \to \infty} x_n = a$, то $\lim\limits_{n \to \infty} |x_n| = |a|$.

Задача 2. Доказать, что последовательность $\{x_n\}$ расходится тогда и только тогда, когда существует такое число $\varepsilon > 0$, что, каково бы ни было действительное число a и каков бы ни был номер n, найдется такой номер $m > n$, для которого выполняется неравенство $|x_m - a| \geqslant \varepsilon$.

УПРАЖНЕНИЕ 3. Записать позитивное определение расходящейся последовательности и условие задачи 2 в логических символах и сравнить их.

В рассмотренных выше примерах существование или отсутствие пределов у данных последовательностей было до-

---

[1] Здесь частица «не» входит не в определение, а в определяемое понятие, что «не является пределом». Само определение начинается со слова «если».

статочно очевидным, а доказательства сводились к элементарной проверке определения предела последовательности.

В качестве более сложного примера отыскания предела последовательности рассмотрим предел средних арифметических членов заданной сходящейся последовательности.

Пример 5. *Если последовательность* $\{x_n\}$ *сходится, то последовательность средних арифметических ее членов*

$$y_n = \frac{x_1 + x_2 + \ldots + x_n}{n}, \quad n = 1, 2, \ldots,$$

также сходится и притом к тому же пределу, что и сама последовательность $\{x_n\}$.

Пусть $\lim\limits_{n \to \infty} x_n = a$. Прежде всего заметим, что для любых натуральных чисел n_0 и $n > n_0$ имеет место равенство

$$y_n - a = \frac{x_1 + \ldots + x_n}{n} - a =$$

$$= \frac{x_1 + \ldots + x_{n_0} - n_0 a}{n} + \frac{(x_{n_0+1} - a) + \ldots + (x_n - a)}{n}. \qquad (4.2)$$

Если теперь задано $\varepsilon > 0$, то, согласно определению предела, существует такой номер n_0, что для всех $n > n_0$ выполняется неравенство

$$|x_n - a| < \frac{\varepsilon}{2}. \qquad (4.3)$$

Число $x_1 + \ldots + x_{n_0} - n_0 a$ фиксировано, а $\lim\limits_{n \to \infty} \frac{1}{n} = 0$, поэтому

$$\lim_{n \to \infty} \frac{x_1 + \ldots + x_{n_0} - n_0 a}{n} = 0.$$

Следовательно, существует такой номер m_0, что для всех $n > m_0$ выполняется неравенство

$$\left|\frac{x_1 + \ldots + x_{n_0} - n_0 a}{n}\right| < \frac{\varepsilon}{2}. \qquad (4.4)$$

Пусть $n_\varepsilon = \max\{n_0, m_0\}$. Тогда для всех номеров $n > n_\varepsilon$ в силу формул (4.2)—(4.4) получим

$$|y_n - a| \leqslant \left|\frac{x_1 + \ldots + x_{n_0} - n_0 a}{n}\right| + \frac{|x_{n_0+1} - a| + \ldots + |x_n - a|}{n} <$$

$$< \frac{\varepsilon}{2} + \frac{n - n_0}{n}\frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Это и означает, что $\lim\limits_{n \to \infty} y_n = a$. □

УПРАЖНЕНИЕ 4. Доказать: 1) что отбрасывание или замена конечного числа элементов последовательности не влияет на ее сходимость, причем в случае сходящейся последовательности не влияет и на величину предела;

2) если $\lim\limits_{n \to \infty} x_n = a$, $\lim\limits_{n \to \infty} y_n = a$ и $z_n = \begin{cases} x_k \text{ при } n = 2k-1, \\ y_k \text{ при } n = 2k, \end{cases}$ $k = 1, 2, \ldots$,

то и $\lim\limits_{n \to \infty} z_n = a$.

Наряду с числовыми последовательностями в данном курсе будут встречаться последовательности точек расширенной числовой прямой, т. е. занумерованные натуральными числами совокупности $\{x_n\}$ элементов расширенного множества действительных чисел $\overline{\boldsymbol{R}}$ (см. п. 3.1). Таким образом, элементами этих последовательностей наряду с действительными числами могут быть бесконечно удаленные точки $+\infty$ и $-\infty$. Для таких последовательностей также можно ввести понятие предела, аналогичное пределу числовых последовательностей и содержащее его в себе как частный случай.

Определение 5. *Точка a расширенной числовой прямой $\overline{\boldsymbol{R}}$ называется пределом последовательности точек этой прямой, если, какова бы ни была окрестность точки a, она содержит все члены рассматриваемой последовательности начиная с некоторого номера.*

Отличие от рассмотренного выше случая состоит в том, что здесь членами последовательности могут быть не только действительные числа, но и бесконечности с определенным знаком.

Конечно, понятие предела можно обобщить и на случай последовательности точек прямой, расширенной с помощью только одной бесконечно удаленной точки — бесконечности без знака.

З а м е ч а н и е. Для любой окрестности $U(a, \varepsilon)$, $\varepsilon > 0$, где a — либо число: $a \in \boldsymbol{R}$, либо одна из бесконечностей $-\infty$, $+\infty$ или ∞, существует такое натуральное n, что выполняется включение $U\left(a, \frac{1}{n}\right) \subset U(a, \varepsilon)$. Чтобы в этом убедиться, достаточно взять такое $n \in \boldsymbol{N}$, что $\frac{1}{n} < \varepsilon$.

Поэтому, если последовательность $x_n \in \boldsymbol{R}$ такова, что для любого $n = 1, 2, \ldots$ выполняется включение $x_n \in U\left(a, \frac{1}{n}\right)$ (здесь a — либо действительное число, либо одна из бесконечностей ∞, $+\infty$ или $-\infty$), то $\lim\limits_{n \to \infty} x_n = a$. В самом деле, для любой окрестности $U(a)$ существует такое натуральное n_0, что $U\left(a, \frac{1}{n_0}\right) \subset U(a)$; тогда для всех номеров $n > n_0$ будем иметь $x_n \in U\left(a, \frac{1}{n}\right) \subset U\left(a, \frac{1}{n_0}\right) \subset U(a)$. Это и означает, что $\lim\limits_{n \to \infty} x_n = a$.

В дальнейшем под последовательностью всегда понимается **числовая последовательность**, т. е. последовательность, элементами которой являются действительные числа, если, конечно, специально не оговорено что-либо другое.

УПРАЖНЕНИЯ. *5.* Привести пример неограниченной последовательности, не являющейся бесконечно большой.

6. Доказать, что если $a_n \leqslant |b_n|$, $n = 1, 2, \ldots$, и $\lim\limits_{n \to \infty} a_n = +\infty$, то $\lim\limits_{n \to \infty} b_n = \infty$.

7. Доказать, что почленное произведение бесконечно большой последовательности на последовательность, абсолютная величина всех членов которой ограничена снизу положительной постоянной, является бесконечно большой последовательностью.

4.2. Единственность предела числовой последовательности

Докажем прежде всего корректность определения предела в том смысле, что если он существует, то он единствен.

ТЕОРЕМА 1. *Последовательность точек расширенной числовой прямой может иметь на этой прямой только один предел.*

СЛЕДСТВИЕ. *Числовая последовательность может иметь только один предел, конечный или определенного знака бесконечный.*

Доказательство теоремы. Допустим, что утверждение теоремы несправедливо. Это означает, что существует последовательность $x_n \in \overline{\boldsymbol{R}}$, $n = 1, 2, \ldots$, у которой имеется

по крайней мере два различных предела: $a \in \overline{\boldsymbol{R}}$ и $b \in \overline{\boldsymbol{R}}$. Выберем $\varepsilon_1 > 0$ и $\varepsilon_2 > 0$ так, чтобы ε_1-окрестность точки a не пересекалась с ε_2-окрестностью точки b. Это всегда можно сделать согласно лемме п. 3.2 (см. рис. 6, *а*, *б*, *в*, *г*). В силу определения предела, из условия $\lim\limits_{n \to \infty} x_n = a$ следует, что существует такой номер $n_1 \in \boldsymbol{N}$, что для всех номеров $n > n_1$, $n \in \boldsymbol{N}$, имеет место включение $x_n \in U(a, \varepsilon_1)$, а из условия $\lim\limits_{n \to \infty} x_n = b$ следует, что существует такое $n_2 \in \boldsymbol{N}$, что для всех $n > n_2$, $n \in \boldsymbol{N}$, справедливо включение $x_n \in U(b, \varepsilon_2)$. Следовательно, если обозначить через n_0 наибольший из номеров n_1 и n_2; $n_0 \overset{\text{def}}{=} \max\{n_1, n_2\}$, то для любого $n > n_0$ одновременно будем иметь $x_n \in U(a, \varepsilon_1)$ и $x_n \in U(b, \varepsilon_2)$, т. е. $x_n \in U(a, \varepsilon_1) \cap U(b, \varepsilon_2)$. Это противоречит условию

$$U(a, \varepsilon_1) \cap U(b, \varepsilon_2) = \varnothing. \ \square$$

Следствие является частным случаем утверждения теоремы.

Для единственности бесконечного предела последовательности элементов из $\overline{\boldsymbol{R}}$ существенным является рассмотрение лишь бесконечностей определенного знака, так как если последовательность имеет своим пределом бесконечность с определенным знаком, то одновременно ее пределом является и бесконечность без знака. Например, если $\lim\limits_{n \to \infty} x_n = +\infty$, то, конечно, и $\lim\limits_{n \to \infty} x_n = \infty$.

Докажем теперь некоторые простые свойства конечных и бесконечных пределов.

4.3. Переход к пределу в неравенствах

Сформулируем и докажем три часто используемых свойства пределов последовательностей точек расширенной числовой прямой, связанные с равенствами и неравенствами для членов последовательностей.

I. *Если для всех* $n = 1, 2, \ldots$ *имеет место равенство* $x_n = a \in \overline{\boldsymbol{R}}$ (*т. е. последовательность* $\{x_n\}$ *стационарна*), *то* $\lim\limits_{n \to \infty} x_n = a$.

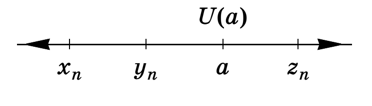


Рис. 10

Коротко говоря, предел постоянной равен самой этой постоянной.

Доказательство. Действительно, в этом случае для любой окрестности $U(a)$ точки a в качестве номера n_0, указанного в определении **5**, можно взять, например, $n_0 = 1$, так как для всех номеров $n = 1, 2, \ldots$ имеет место включение

$$x_n = a \in U(a). \ \square$$

II. *Если*

$$x_n \in \overline{\boldsymbol{R}}, \quad y_n \in \overline{\boldsymbol{R}}, \quad z_n \in \overline{\boldsymbol{R}}, \quad x_n \leqslant y_n \leqslant z_n, \quad n = 1, 2, \ldots, \tag{4.5}$$

и

$$\lim_{n \to \infty} x_n = \lim_{n \to \infty} z_n = a \in \overline{\boldsymbol{R}}, \tag{4.6}$$

то

$$\lim_{n \to \infty} y_n = a. \tag{4.7}$$

Доказательство. Зафиксируем произвольно окрестность $U(a)$ точки a. В силу условий (4.6), существуют такой номер n_1, что для всех номеров $n > n_1$ выполняется включение

$$x_n \in U(a), \tag{4.8}$$

и такой номер n_2, что для всех номеров $n > n_2$ — включение

$$z_n \in U(a). \tag{4.9}$$

Возьмем в качестве номера n_0 наибольший из номеров n_1 и n_2: $n_0 = \max\{n_1;\ n_2\}$. Тогда для номеров $n > n_0$ одновременно выполняются включения (4.8) и (4.9); тогда (рис. 10), в силу условия (4.5), для всех $n > n_0$ выполняется включение $y_n \in U(a)$, а это и означает справедливость утверждения (4.7). $\square$

СЛЕДСТВИЕ. *Если* $x_n \leqslant y_n$, $x_n \in \overline{\boldsymbol{R}}$, $y_n \in \overline{\boldsymbol{R}}$, $n = 1, 2, \ldots$, *и*

$$\lim_{n \to \infty} x_n = +\infty, \tag{4.10}$$

то

$$\lim_{n \to \infty} y_n = +\infty, \tag{4.11}$$

а если $\lim\limits_{n \to \infty} y_n = -\infty$, *то* $\lim\limits_{n \to \infty} x_n = -\infty$.

Доказательство. Пусть выполнено условие (4.10). Рассмотрим вспомогательную последовательность $z_n = +\infty$, $n = 1, 2, \ldots$, тогда, очевидно, для последовательностей $\{x_n\}$, $\{y_n\}$, $\{z_n\}$ выполняются условия (4.5) и (4.6) при $a = +\infty$, а поэтому, в силу утверждения (4.7), имеет место и равенство (4.11).

Рис. 11

Аналогично рассматривается и случай $\lim\limits_{n \to \infty} y_n = -\infty$. □

III. *Если* $x_n \in \overline{\boldsymbol{R}}$, $y_n \in \overline{\boldsymbol{R}}$, $n = 1, 2, \ldots$, *и*

$$\lim_{n \to \infty} x_n = a, \quad \lim_{n \to \infty} y_n = b, \tag{4.12}$$

причем

$$a < b,\ a \in \overline{\boldsymbol{R}},\ b \in \overline{\boldsymbol{R}}, \tag{4.13}$$

то существует такой номер n_0, *что для всех номеров* $n > n_0$ *выполняется неравенство*

$$x_n < y_n. \tag{4.14}$$

Доказательство. Пусть $U = U(a)$ и $V = V(b)$ — какие-либо непересекающиеся окрестности точек a и b, тогда из условия $a < b$ следует, что для любых $x \in U$ и $y \in V$ выполняется неравенство (рис. 11)

$$x < y. \tag{4.15}$$

В силу условия (4.12), существует такой номер n_0, что для всех номеров $n > n_0$ выполняются включения

$$x_n \in U,\ y_n \in V, \tag{4.16}$$

а поэтому, согласно неравенству (4.15), имеют место неравенства (4.14). □

СЛЕДСТВИЕ 1. *Пусть* a, b, x_n *принадлежат* $\overline{\boldsymbol{R}}$, $n = 1, 2, \ldots$. *Если* $\lim\limits_{n \to \infty} x_n = a$ *и* $a < b$ (*соответственно* $a > b$), *то существует такой номер* n_0, *что для всех номеров* $n > n_0$ *выполняется неравенство*

$$x_n < b \tag{4.17}$$

(*соответственно неравенство* $x_n > b$).

Доказательство. Пусть $a < b$. Рассмотрим вспомогательную последовательность $y_n = b$, $n = 1, 2, \ldots$; тогда для последовательностей $\{x_n\}$ и $\{y_n\}$ выполняются условия (4.12) и (4.13), следовательно, и условие (4.14), которое в данном случае превращается в неравенство (4.17).

Аналогично рассматривается случай $a > b$. □

СЛЕДСТВИЕ 2. *Если* $\lim\limits_{n \to \infty} x_n = a$, $\lim\limits_{n \to \infty} y_n = b$, $x_n \in \overline{\boldsymbol{R}}$, $y_n \in \overline{\boldsymbol{R}}$, $n = 1, 2, \ldots$, $a \in \overline{\boldsymbol{R}}$, $b \in \overline{\boldsymbol{R}}$ *и для всех* $n = 1, 2, \ldots$ *выполняется неравенство*

$$x_n \leqslant y_n, \tag{4.18}$$

то

$$a \leqslant b. \tag{4.19}$$

Доказательство. Пусть выполнено условие (4.18). Если бы оказалось, что $a > b$, то, согласно свойству III пределов, нашелся бы такой номер n_0, что для всех номеров $n > n_0$ выполнялось бы неравенство $x_n > y_n$, что противоречит условию (4.18). Следовательно, выполняется неравенство (4.19). □

Из следствия 2 вытекает, в частности, что если $x_n \leqslant b$, $n = 1, 2, \ldots$, и $\lim\limits_{n \to \infty} x_n = a$, то имеет место неравенство $a \leqslant b$.

В самом деле, если взять вспомогательную стационарную последовательность $y_n = b$, $n = 1, 2, \ldots$, то для последовательностей $\{x_n\}$ и $\{y_n\}$ будут выполняться условия следствия 2, т. е. $\lim\limits_{n \to \infty} x_n = a$, $x_n \in \overline{\boldsymbol{R}}$, $n = 1, 2, \ldots$, $a \in \overline{\boldsymbol{R}}$, и для всех $n = 1, 2, \ldots$ справедливы неравенства $x_n \leqslant y_n = b$. Поэтому, согласно следствию 2, имеет место и неравенство $a \leqslant b$. □

Следствие 2 означает, что если последовательности $\{x_n\}$ и $\{y_n\}$ имеют пределы $\lim\limits_{n \to \infty} x_n = a$, $\lim\limits_{n \to \infty} y_n = b$, $a \in \overline{\boldsymbol{R}}$, $b \in \overline{\boldsymbol{R}}$, то в неравенствах $x_n < y_n$ и $x_n \leqslant y_n$ можно переходить к пределу, причем даже в первом случае в результате получается, вообще говоря, нестрогое неравенство

$$a = \lim_{n \to \infty} x_n \leqslant \lim_{n \to \infty} y_n = b.$$

Отметим, что нас в основном интересуют числовые последовательности. Последовательности же точек расширенной числовой прямой введены прежде всего для большей компактности изложения: они позволяют не рассматривать отдельно случаи конечных и определенного знака бесконечных пределов последовательностей. В дальнейшем определения и утверждения будут в основном формулироваться для числовых последовательностей, хотя многие из них без всякого труда обобщаются на случай последовательностей точек расширенной числовой прямой.

З а м е ч а н и е 1. Для того чтобы числовая последовательность $\{x_n\}$ имела своим пределом число a, необходимо, чтобы для любых двух чисел $\varepsilon > 0$ и $c > 0$ существовал такой номер n_0, что для всех номеров $n > n_0$ выполняется неравенство $|x_n - a| < c\varepsilon$ и достаточно, чтобы нашлось такое число $c > 0$, что для любого $\varepsilon > 0$ существовал такой номер n_0, что для всех номеров $n > n_0$ выполнялось то же неравенство $|x_n - a| < c\varepsilon$.

Оба этих утверждения следуют из того, что при фиксированном $c > 0$ при произвольном $\varepsilon > 0$ число $c\varepsilon$ также является произвольным положительным числом.

Например, если $\lim\limits_{n \to \infty} x_n = a$, то для всякого $\varepsilon > 0$ существует такой номер n_0, что для всех номеров $n > n_0$ выполняется неравенство $|x_n - a| < \varepsilon/2$.

Иногда бывает полезно рассмотреть последовательность, получающуюся из данной последовательности перенумерацией всех ее членов или некоторого бесконечного подмножества. В дальнейшем для таких последовательностей будет неоднократно использоваться следующая лемма.

ЛЕММА. *Если последовательность $x_n \in \boldsymbol{R}$, $n = 1, 2, \ldots$, имеет конечный или бесконечный предел и $\{n_k\}$ — такая последовательность натуральных чисел, что*

$$\lim_{k \to \infty} n_k = +\infty, \tag{4.20}$$

то последовательность $\{x_{n_k}\}$ имеет тот же предел, что и последовательность $\{x_n\}$.

В обозначении x_{n_k} члена последовательности $\{x_{n_k}\}$ номер n_k является номером этого члена в последовательности $\{x_n\}$, а номер k — его номером в последовательности $\{x_{n_k}\}$.

Доказательство. Пусть $\lim\limits_{n\to\infty} x_n = a$. Это означает, что для любой окрестности $U(a)$ точки a существует такой номер n_0, что для всех номеров $n > n_0$ выполняется включение $x_n \in \in U(a)$. В силу выполнения условия (4.20), для номера n_0 существует такой номер k_0, что для всех номеров $k > k_0$ справедливы неравенства $n_k > n_0$ и, следовательно, имеет место включение $x_{n_k} \in U(a)$. Это и означает, что $\lim\limits_{k\to\infty} x_{n_k} = a$. □

Замечание 2. Если $\lim\limits_{n\to\infty} x_n = \infty$ и $\lim\limits_{k\to\infty} n_k = +\infty$, то иногда о пределе последовательности $\{x_{n_k}\}$ можно сказать больше, чем только то, что $\lim\limits_{k\to\infty} x_{n_k} = \infty$: может оказаться, что $\lim\limits_{k\to\infty} x_{n_k} = +\infty$ или $\lim\limits_{k\to\infty} x_{n_k} = -\infty$. Например,

$$\lim_{n\to\infty} (-n)^n = \infty, \text{ а } \lim_{k\to\infty} (-2k)^{2k} = +\infty,$$
$$\lim_{k\to\infty} [-(2k-1)]^{2k-1} = -\infty.$$

Определение 6. *Последовательность $\{x_{n_k}\}$, которая составлена из членов последовательности $\{x_n\}$ и в которой порядок следования ее элементов совпадает с их порядком следования в исходной последовательности $\{x_n\}$, называется подпоследовательностью этой последовательности.*

Таким образом, последовательность $\{x_{n_k}\}$ является подпоследовательностью последовательности $\{x_n\}$, если условие $k < k'$ равносильно условию $n_k < n_{k'}$, $k, k' = 1, 2, \ldots$.

Так, последовательность $1, 3, 5, \ldots, 2n+1, \ldots$ является, а последовательность $2, 1, 3, 4, \ldots, n, \ldots$ не является подпоследовательностью натурального ряда чисел $1, 2, \ldots, n, \ldots$. В обоих случаях элементы последовательностей образуют подмножество[1] множества натуральных чисел, но в первом

[1] Напомним (см. п. 1.1), что само множество также считается своим подмножеством.

случае члены последовательности расположены в том же порядке, что и в натуральном ряде чисел, а во втором случае этот порядок нарушен.

Если $\{x_{n_k}\}$ — подпоследовательность последовательности $\{x_n\}$, то, очевидно, $n_k \geqslant k$, $k = 1, 2, \ldots$, и, следовательно, $\lim\limits_{k \to \infty} n_k = +\infty$. Отсюда следует, в силу леммы, что если последовательность имеет конечный или бесконечный предел, то и любая ее подпоследовательность имеет тот же предел.

УПРАЖНЕНИЕ 8. Пусть из членов последовательности натуральных чисел $\{n\}$ образована новая последовательность n_k так, что разные члены последовательности $\{n\}$ переходят в разные члены новой и каждый член последовательности $\{n\}$ переходит в некоторый член новой. Это равносильно тому, что задана биекция $n \mapsto n_k$ множества натуральных чисел $\boldsymbol{N}$ на себя. Доказать, что в этом случае $\lim\limits_{k \to \infty} n_k = \infty$.

4.4. Ограниченность сходящихся последовательностей

Следует различать *последовательность* $\{x_n\}$, т. е. множество элементов x_n и *множество значений ее элементов*. Первое множество всегда бесконечно, так как состоит из совокупности элементов, отличающихся по крайней мере номерами $n = 1, 2, \ldots$. Второе множество состоит из всех чисел, являющихся значениями элементов данной последовательности, оно может быть и конечным. Например, стационарная последовательность $x_n = 1$, $n = 1, 2, \ldots$, как и всякая последовательность, состоит из бесконечного числа элементов, а множество значений ее элементов состоит из одного числа 1.

Определение 7. *Последовательность называется ограниченной сверху* (*снизу*), *если множество значений ее элементов ограничено сверху* (*снизу*).

В терминах элементов последовательности это определение может быть перефразировано следующим образом.

Определение 7′. *Последовательность* $\{x_n\}$ *называется ограниченной сверху* (*снизу*), *если существует такое число* b,

что для всех номеров $n = 1, 2, \ldots$ выполняется неравенство $x_n \leqslant b$ (соответственно неравенство $x_n \geqslant b$).

Определение 8. *Последовательность, ограниченная сверху и снизу, называется просто ограниченной.*

Очевидно, что последовательность $\{x_n\}$ ограничена тогда и только тогда, когда существует такое число b, что для всех номеров $n = 1, 2, \ldots$ выполняется неравенство $|x_n| \leqslant b$.

Определение 9. *Последовательность, не являющаяся ограниченной (сверху, снизу), называется неограниченной (сверху, снизу).*

Например, последовательности $\left\{\frac{1}{n}\right\}$ и $\left\{\sin\frac{\pi}{2}n\right\}$ ограничены. Последовательность $\{n\}$ не ограничена, точнее, она ограничена снизу, но не ограничена сверху, а последовательность $\left\{n\sin\frac{\pi}{2}n\right\}$ является неограниченной как сверху, так и снизу.

ТЕОРЕМА 2. *Если числовая последовательность имеет конечный предел, то она ограничена.*

Доказательство. Пусть дана сходящаяся последовательность $\{x_n\}$ и пусть $\lim\limits_{n\to\infty} x_n = a$. Возьмем, например, $\varepsilon = 1$. Согласно определению предела последовательности, существует такое n_1, что для всех $n > n_1$ выполняется неравенство $|x_n - a| < 1$. Пусть d — наибольшее из чисел $1, |x_1 - a|, \ldots$ $\ldots, |x_{n_1} - a|$. Тогда для всех $n = 1, 2, \ldots$ справедливо неравенство $|x_n - a| \leqslant d$, т. е. для всех n

$$a - d \leqslant x_n \leqslant a + d.$$

Это и означает ограниченность заданной последовательности. □

4.5. Монотонные последовательности

Определение 10. *Верхняя (нижняя) грань множества значений элементов последовательности $\{x_n\}$ называется верхней (нижней) гранью данной последовательности и*

обозначается $\sup\{x_n\}$ *или* $\sup\limits_{n=1,2,\ldots} x_n$ (*соответственно* $\inf\{x_n\}$ *или* $\inf\limits_{n=1,2,\ldots} x_n$).

Если верхняя (нижняя) грань является числом, то это определение можно сформулировать следующим образом.

Определение 11. *Число a является верхней* (*нижней*) *гранью последовательности x_n, $n = 1, 2, \ldots$, если*:

1) *для всех $n = 1, 2, \ldots$ выполняется неравенство $x_n \leqslant a$* (*неравенство $x_n \geqslant a$*);

2) *для любого $\varepsilon > 0$ существует такой номер n_ε, что* $x_{n_\varepsilon} > a - \varepsilon$ (соответственно $x_{n_\varepsilon} < a + \varepsilon$).

Аналогично можно сформулировать определение верхней (нижней) грани последовательности в том случае, когда указанная грань бесконечна. (Сделайте это.)

В качестве примеров отметим, что $\sup\{1/n\} = 1$, $\inf\{1/n\} = 0$, $\sup\{n\} = +\infty$, $\inf\{n\} = 1$. Здесь везде $n = 1, 2, \ldots$.

Определение 12. *Последовательность $\{x_n\}$ называется возрастающей* (*убывающей*) *последовательностью, если для каждого $n = 1, 2, \ldots$ выполняется неравенство $x_n \leqslant x_{n+1}$* (*соответственно неравенство $x_n \geqslant x_{n+1}$*)[1].

Возрастающие и убывающие последовательности называются *монотонными*.

Например, последовательность $\left\{\frac{1}{n}\right\}$ убывает, последовательность $\{n\}$ возрастает, а последовательность $\left\{\sin\frac{\pi}{2}n\right\}$ не является монотонной.

ТЕОРЕМА 3 (теорема Вейерштрасса[2]). *Всякая возрастающая числовая последовательность $\{x_n\}$ имеет предел, конечный, если она ограничена сверху, и бесконечный, равный $+\infty$, если она не ограничена сверху, причем*

$$\lim_{n\to\infty} x_n = \sup\{x_n\}. \qquad (4.21)$$

---

[1] Возрастающие (убывающие) последовательности называются также неубывающими (невозрастающими).

[2] К. Вейерштрасс (1815—1897) — немецкий математик.

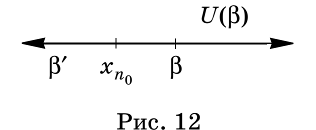


Рис. 12

Аналогично, всякая убывающая числовая последовательность $\{x_n\}$ имеет предел, конечный, если эта последовательность ограничена снизу, и бесконечный, равный $-\infty$, если она не ограничена снизу, причем

$$\lim_{n \to \infty} x_n = \inf \{x_n\}. \tag{4.22}$$

Доказательство. Пусть последовательность $\{x_n\}$ возрастает. Докажем равенство (4.21). Остальные утверждения теоремы для возрастающих последовательностей следуют из него очевидным образом.

Пусть $\beta = \sup \{x_n\}$ (значение β может быть как конечным, так и бесконечным, равным $+\infty$). Возьмем произвольную окрестность $U(\beta)$ точки β и обозначим через β' ее левый конец (рис. 12). Очевидно, $\beta' < \beta$.

Согласно определению верхней грани:

1) для любого номера $n \in \boldsymbol{N}$ имеет место неравенство

$$x_n \leqslant \beta; \tag{4.23}$$

2) существует такой номер n_0, что

$$x_{n_0} > \beta'. \tag{4.24}$$

В силу возрастания последовательности $\{x_n\}$, из (4.23) и (4.24) следует, что для всех номеров $n > n_0$ выполняется неравенство

$$\beta' \underset{(4.24)}{<} x_{n_0} \leqslant x_n \underset{(4.23)}{\leqslant} \beta, \tag{4.25}$$

поэтому при $n > n_0$ имеет место включение

$$x_n \underset{(4.25)}{\in} U(\beta),$$

а это и означает, что β является пределом последовательности $\{x_n\}$.

Аналогично рассматривается случай убывающей последовательности. □

Замечание 1. Таким образом, всякая монотонная последовательность имеет предел: конечный, если она ограничена, и бесконечный, если она не ограничена. Этот предел равен $+\infty$, если монотонная последовательность не ограничена сверху, и равен $-\infty$, если она не ограничена снизу.

Всякая подпоследовательность монотонной последовательности также монотонна, поэтому она, в свою очередь, всегда имеет конечный или бесконечный предел, который, очевидно, совпадает с пределом всей последовательности (см. лемму в п. 4.3).

Было показано, что если последовательность сходится, то она ограничена (теорема 2), отсюда, в частности, следует, что если возрастающая последовательность сходится, то она ограничена сверху; с другой стороны, если возрастающая последовательность ограничена сверху, то она сходится (теорема 3). Таким образом, справедливо следующее утверждение.

СЛЕДСТВИЕ. *Для того чтобы возрастающая последовательность сходилась, необходимо и достаточно, чтобы она была ограничена сверху.*

Аналогичное утверждение справедливо и для убывающей последовательности.

Замечание 2. Если $\{[a_n, b_n]\}$ — система вложенных отрезков, по длине стремящихся к нулю, а ξ — точка, принадлежащая всем отрезкам данной системы, то

$$\xi = \lim_{n \to \infty} a_n = \lim_{b \to \infty} b_n. \tag{4.26}$$

В самом деле, в п. 3.7 было показано, что $\xi = \sup \{a_n\} = \inf \{b_n\}$. С другой стороны, последовательность $\{a_n\}$, соответственно $\{b_n\}$ возрастает (убывает), откуда в силу равенств (4.21) и (4.22) и следует (4.26).

Пример. Число e.

Пусть $x_n = \left(1 + \frac{1}{n}\right)^n$, $n = 1, 2, \ldots$.

Покажем, что эта последовательность сходится. Применяя формулу бинома Ньютона, получаем

$$x_n = \left(1 + \frac{1}{n}\right)^n = 1 + n \cdot \frac{1}{n} + \frac{n(n-1)}{1 \cdot 2} \cdot \frac{1}{n^2} + \frac{n(n-1)(n-2)}{1 \cdot 2 \cdot 3} \cdot \frac{1}{n^3} + \ldots$$

$$\ldots + \frac{n(n-1)\ldots(n-k+1)}{1 \cdot 2 \ldots k} \cdot \frac{1}{n^k} + \ldots + \frac{n(n-1)\ldots 1}{1 \cdot 2 \ldots n} \cdot \frac{1}{n^n} =$$

$$= 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n}\right) + \frac{1}{3!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) + \ldots +$$

$$+ \frac{1}{k!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ldots\left(1 - \frac{k-1}{n}\right) + \ldots + \frac{1}{n!}\left(1 - \frac{1}{n}\right)\ldots\left(1 - \frac{n-1}{n}\right). \tag{4.27}$$

При переходе от n к $n + 1$ в сумме, стоящей в правой части равенства (4.27), число слагаемых, которые все положительны, возрастает и, кроме того, каждое слагаемое начиная с третьего увеличивается, так как

$$1 - \frac{s}{n} < 1 - \frac{s}{n+1}, \ s = 1, 2, \ldots, n - 1, n = 2, 3, \ldots,$$

поэтому

$$x_n < x_{n+1}, \ n = 1, 2, \ldots .$$

Далее, замечая, что в (4.27) каждая из скобок вида $\left(1 - \frac{s}{n}\right)$ меньше единицы и что

$$2^{n-1} = \underbrace{1 \cdot 2 \cdot 2 \cdot \ldots \cdot 2}_{n \text{ сомножителей}} \leqslant 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n = n!,$$

следовательно, $\frac{1}{n!} \leqslant \frac{1}{2^{n-1}}$, для всех $n = 1, 2, 3, \ldots$, имеем

$$x_n < 2 + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{n!} \leqslant 2 + \frac{1}{2} + \frac{1}{4} + \ldots + \frac{1}{2^{n-1}}. \quad (4.28)$$

Сумма $\frac{1}{2} + \frac{1}{2^2} + \ldots + \frac{1}{2^{n-1}}$ (которую легко подсчитать по известной из элементарной математики формуле для суммы членов геометрической прогрессии: она равна $1 - \frac{1}{2^{n-1}}$) при любом $n = 1, 2, \ldots$ меньше единицы, поэтому окончательно получим

$$2 \leqslant x_n < x_{n+1} < 3. \quad (4.29)$$

Итак, последовательность $\{x_n\}$ возрастает и ограничена сверху, а поэтому, согласно теореме 3, имеет предел. Этот предел и обозначается буквой e.

Переходя к пределу в (4.29), получим $2 < e \leqslant 3$. Более точными оценками можно получить, что справедливо приближенное равенство

$$e \approx 2{,}718281828459045.$$

Доказывается также, что число e иррационально (см. п. 34.14*) и, более того, трансцендентно, т. е. не является корнем никакого алгебраического уравнения с целыми ко-

эффициентами. Число e в математическом анализе играет особую роль. Оно, в частности, является основанием натуральных логарифмов.

4.6. Теорема Больцано—Вейерштрасса

В п. 4.4 было доказано, что всякая сходящаяся последовательность ограничена. Обратное утверждение, конечно, неверно. Например, последовательность $x_n = (-1)^n$, $n = 1, 2, \ldots$, ограничена и расходится. Однако оказывается, что всякая ограниченная последовательность содержит сходящуюся подпоследовательность. Это утверждение называется теоремой Больцано—Вейерштрасса[1] или свойством компактности ограниченной последовательности.

ТЕОРЕМА 4. *Из любой ограниченной последовательности можно выделить сходящуюся подпоследовательность, а из любой неограниченной последовательности — бесконечно большую подпоследовательность, имеющую своим пределом бесконечность определенного знака.*

Доказательство. Пусть последовательность $\{x_n\}$ ограничена, т. е. существует такой отрезок $[a, b]$, что $a \leqslant x_n \leqslant b$ для всех $n = 1, 2, \ldots$. Разделим отрезок $[a, b]$ на два равных отрезка. По крайней мере один из получившихся отрезков содержит бесконечно много элементов данной последовательности. Обозначим его через $[a_1, b_1]$. Пусть x_{n_1} — какой-либо из членов данной последовательности, лежащий на отрезке $[a_1, b_1]$.

Разделим отрезок $[a_1, b_1]$ на два равных отрезка; снова хотя бы один из получившихся отрезков содержит бесконечно много членов исходной последовательности; обозначим его через $[a_2, b_2]$. В силу того что на отрезке $[a_2, b_2]$ бесконечно много членов последовательности $\{x_n\}$, найдется такой ее член x_{n_2}, что $x_{n_2} \in [a_2, b_2]$ и $n_2 > n_1$. Продолжая этот процесс, получим последовательность отрезков $[a_k, b_k]$, в которой каждый последующий является половиной предыдущего, и последовательность таких элементов x_{n_k} данной после-

[1] Б. Больцано (1781—1848) — чешский математик.

довательности, что $x_{n_k} \in [a_k, b_k]$, $k = 1, 2, \ldots$, и $n_{k''} > n_{k'}$ при $k'' > k'$. Последовательность $\{x_{n_k}\}$ является, в силу построения, подпоследовательностью последовательности $\{x_n\}$. Покажем, что эта подпоследовательность сходящаяся.

Последовательность отрезков $[a_k, b_k]$, $k = 1, 2, \ldots$, является последовательностью вложенных отрезков, по длине стремящихся к нулю, так как $b_k - a_k = \frac{b - a}{2^k} \to 0$ при $k \to \infty$. Согласно принципу вложенных отрезков (см. п. 3.7), существует единственная точка ξ, принадлежащая всем этим отрезкам. Как было показано (см. формулу (4.26)) в замечании 2 к теореме 3, $\lim\limits_{k \to \infty} a_k = \lim\limits_{k \to \infty} b_k = \xi$, но $a_k \leqslant x_{n_k} \leqslant b_k$, $k = 1, 2, \ldots$, поэтому, согласно свойству II (см. п. 4.3 сходящихся последовательностей), последовательность $\{x_{n_k}\}$ также сходится и $\lim\limits_{k \to \infty} x_{n_k} = \xi$.

Пусть теперь последовательность $\{x_n\}$ не ограничена. Тогда она либо не ограничена сверху, либо не ограничена снизу, либо имеет место и то и другое. Пусть для определенности последовательность $\{x_n\}$ не ограничена сверху. Тогда существует такой номер $n_1 \in \boldsymbol{N}$, что $x_{n_1} > 1$.

Очевидно, последовательность x_n, $n = n_1 + 1, n_1 + 2, \ldots$, также не ограничена сверху, так как получается из данной неограниченной сверху последовательности x_n, $n = 1, 2, \ldots$, отбрасыванием конечного числа ее членов. Поэтому существует такое $n_2 > n_1$, $n_2 \in \boldsymbol{N}$, что $x_{n_2} > 2$.

Продолжая этот процесс, получаем последовательность таких номеров n_k, что $n_1 < n_2 < \ldots < n_k < \ldots$ и $x_{n_1} > 1$, $x_{n_2} > 2, \ldots x_{n_k} > k \ldots$. Отсюда следует, что $\{x_{n_k}\}$ — подпоследовательность последовательности $\{x_n\}$ и, согласно следствию свойства II п. 4.3, что $\lim\limits_{k \to \infty} x_{n_k} = +\infty$. □

З а м е ч а н и е. Второе утверждение теоремы 4 можно уточнить. В доказательстве теоремы 4 было показано, что если последовательность не ограничена сверху, то у нее су-

ществует подпоследовательность, стремящаяся к $+\infty$. Аналогично, если последовательность не ограничена снизу, то у нее существует подпоследовательность, стремящаяся к $-\infty$.

Определение 13. *Предел, конечный или определенного знака бесконечный, подпоследовательности данной последовательности называется ее частичным пределом.*

Теорема Больцано—Вейерштрасса (первая часть теоремы 4) и ее аналог для неограниченных последовательностей (вторая часть теоремы 4) показывают, что

всякая последовательность имеет хотя бы один частичный конечный или бесконечный предел, причем заведомо конечный, если данная последовательность ограничена.

Таким образом, каждая числовая последовательность $\{x_n\}$, $x_n \in \boldsymbol{R}$, имеет хотя бы один частичный предел в расширенном множестве действительных чисел, т. е. множество частичных пределов в $\overline{\boldsymbol{R}}$ для любой последовательности всегда не пусто.

УПРАЖНЕНИЯ. 9. Доказать, что, для того чтобы последовательность была сходящейся, необходимо и достаточно, чтобы она была ограничена и имела единственный частичный предел.

10. Доказать, что элемент a (число или одна из бесконечностей со знаком: $+\infty$ или $-\infty$) является частичным пределом последовательности тогда и только тогда, когда в любой его окрестности содержится бесконечно много членов данной последовательности.

4.7. Критерий Коши сходимости последовательности

До сих пор не было дано достаточно общего критерия, с помощью которого можно было бы узнать, сходится ли данная последовательность. Само определение сходящейся последовательности для этого неудобно, так как в него входит значение предела, которое может быть и неизвестным. Поэтому желательно иметь такой критерий для определения сходимости и расходимости последовательностей, который основывался бы только на свойствах элементов данной последовательности. Следующая ниже теорема 5 и дает как раз подобный критерий.

Определение 14. *Будем говорить, что последовательность $\{x_n\}$ удовлетворяет условию Коши*[1]*, если для любого $\varepsilon > 0$ существует такой номер n_ε, что для всех номеров n и m, удовлетворяющих условию $n > n_\varepsilon$, $m > n_\varepsilon$, справедливо неравенство*

$$|x_n - x_m| < \varepsilon. \qquad (4.30)$$

Последовательности, удовлетворяющие условию Коши, называются также *фундаментальными последовательностями*.

С помощью логических символов условие Коши записывается следующим образом:

$$\forall\, \varepsilon > 0\ \exists\, n_\varepsilon \in \boldsymbol{N}\ \forall\, n \in \boldsymbol{N}\ \forall\, m \in \boldsymbol{N},\ n > n_\varepsilon,\ m > n_\varepsilon : |x_n - x_m| < \varepsilon.$$

Условие (4.30) можно сформулировать и так:

для любого $\varepsilon > 0$ существует такой номер n_ε, что для всех номеров $n > n_\varepsilon$ и всех целых неотрицательных p

$$|x_{n+p} - x_n| < \varepsilon. \qquad (4.31)$$

Для того чтобы убедиться в равносильности условий (4.30) и (4.31), достаточно положить $p = n - m$, если $n \geqslant m$, и $p = m - n$, если $m > n$.

ТЕОРЕМА 5 (критерий Коши). *Для того чтобы последовательность сходилась, необходимо и достаточно, чтобы она удовлетворяла условию Коши.*

Доказательство необходимости. Пусть последовательность $\{x_n\}$ сходится и $\lim\limits_{n \to \infty} x_n = a$. Зададим $\varepsilon > 0$; тогда, согласно определению предела последовательности, существует такое n_ε, что для всех номеров $n > n_\varepsilon$ выполняется неравенство $|x_n - a| < \frac{\varepsilon}{2}$.

Пусть теперь $n > n_\varepsilon$ и $m > n_\varepsilon$; тогда

$$|x - x_m| = |(x_n - a) + (a - x_m)| \leqslant |x_n - a| + |x_m - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

т. е. выполняется условие Коши.

Доказательство достаточности. Пусть последовательность удовлетворяет условию Коши, т. е. для всякого $\varepsilon > 0$ существует такое n_ε, что если $n > n_\varepsilon$ и $m > n_\varepsilon$, то

---

[1] О. Коши (1798—1857) — французский математик.

$|x_n - x_m| < \varepsilon$. Возьмем, например, $\varepsilon = 1$; тогда существует такое n_1, что при $n > n_1$ и $m > n_1$ выполняется неравенство $|x_n - x_m| < 1$. В частности, если $n > n_1$ и $m = n_1 + 1$, то $|x_{n_1} - x_{n+1}| < 1$, т. е. $x_{n_1+1} - 1 < x_n < x_{n_1+1} + 1$ при $n > n_1$. Это значит, что последовательность x_n, $n = n_1 + 1, n_1 + 2, \ldots$, ограничена. Поэтому, в силу теоремы 4, существует ее сходящаяся подпоследовательность $\{x_{n_k}\}$.

Пусть $\lim\limits_{k \to \infty} x_{n_k} = a$. Покажем, что вся данная последовательность $\{x_n\}$ также сходится и имеет пределом число a. Зададим некоторое $\varepsilon > 0$. Тогда, во-первых, по определению предела последовательности, существует такое k_ε, что для всех номеров $k > k_\varepsilon$ или, что то же самое, согласно определению подпоследовательности, для всех $n_k > n_{k_\varepsilon}$ выполняется неравенство $|x_{n_k} - a| < \frac{\varepsilon}{2}$.

Во-вторых, так как последовательность $\{x_n\}$ удовлетворяет условию Коши, то существует такое n_ε, что для всех $n > n_\varepsilon$ и всех $m > n_\varepsilon$ выполняются неравенства $|x_n - x_m| < \frac{\varepsilon}{2}$.

Положим $N_\varepsilon = \max\{n_\varepsilon, n_{k_\varepsilon}\}$ и зафиксируем некоторое $n_k > N_\varepsilon$. Тогда для всех $n > N_\varepsilon$ получим

$$|x_n - a| = |(x_n - x_{n_k}) + (x_{n_k} - a)| \leqslant$$

$$\leqslant |x_n - x_{n_k}| + |x_{n_k} - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

а это и доказывает, что $\lim\limits_{n \to \infty} x_n = a$. □

УПРАЖНЕНИЯ. 11. Сформулировать позитивные необходимые и достаточные условия, являющиеся отрицанием критерия Коши, для того чтобы последовательность не имела предела.

12. Доказать, что, для того чтобы последовательность $\{x_n\}$ была сходящейся, необходимо и достаточно, чтобы для любого $\varepsilon > 0$ существовало такое $n_\varepsilon \in \mathbf{N}$, что для всех $n > n_\varepsilon$ выполнялось неравенство $|x_n - x_{n_\varepsilon}| < \varepsilon$.

Задача 3. Выяснить, вытекает или нет сходимость последовательности $\{x_n\}$ из условия, что для любого натурального p существует предел $\lim\limits_{n \to \infty} (x_{n+p} - x_n) = 0$.

4.8. Бесконечно малые последовательности

Над последовательностями можно производить арифметические операции сложения, вычитания, умножения и деления. Определим их.

Определение 15. *Пусть заданы числовые последовательности* $\{x_n\}$ *и* $\{y_n\}$: *суммой, разностью и произведением этих последовательностей называются соответственно последовательности* $\{x_n + y_n\}$, $\{x_n - y_n\}$ *и* $\{x_n y_n\}$. *Если* $y_n \neq 0$, $n = 1, 2, \ldots$, *то частным от деления последовательности* $\{x_n\}$ *на последовательность* $\{y_n\}$ *называется последовательность* $\{x_n/y_n\}$. *Наконец, произведением последовательности* $\{x_n\}$ *на число c называется последовательность* $\{cx_n\}$.

Если последовательность $\{y_n\}$ такова, что в ней имеется лишь конечное число элементов, равных нулю, т. е. существует такое $n_0 \in \boldsymbol{N}$, что при $n \geqslant n_0$, $n \in \boldsymbol{N}$, выполняется неравенство $y_n \neq 0$, то можно рассматривать последовательность $\{x_n/y_n\}$, понимая под ней последовательность с номерами $n \geqslant n_0$.

Определение 16. *Последовательность* $\{\alpha_n\}$ *называется бесконечно малой последовательностью, если* $\lim\limits_{n \to \infty} \alpha_n = 0$.

В п. 3.1 уже были рассмотрены бесконечно малые последовательности $\alpha_n = \frac{1}{n}$, $\alpha_n = \frac{1}{n} \sin \frac{\pi}{2} n$, $n = 1, 2, \ldots$.

Отметим несколько свойств бесконечно малых числовых последовательностей.

I. *Любая конечная линейная комбинация бесконечно малых является бесконечно малой.*

Доказательство. Пусть числовые последовательности $\{\alpha_n\}$ и $\{\beta_n\}$ — бесконечно малые, т. е.

$$\lim_{n \to \infty} \alpha_n = \lim_{\beta \to \infty} \beta_n = 0, \qquad (4.32)$$

а λ и μ — какие-либо действительные числа. Покажем, что последовательность $\{\lambda\alpha_n + \mu\beta_n\}$ также бесконечно малая. Зададим произвольно $\varepsilon > 0$ и возьмем какое-либо число c такое, что

$$c > |\lambda| + |\mu|. \tag{4.33}$$

Тогда, согласно определению предела, из (4.32) следует, что существует такой номер n_0, что для всех номеров $n > n_0$ выполняются неравенства

$$|\alpha_n| < \frac{\varepsilon}{c}, \quad |\beta_n| < \frac{\varepsilon}{c}, \tag{4.34}$$

следовательно, и неравенство

$$|\lambda\alpha_n + \mu\beta_n| \leqslant |\lambda||\alpha_n| + |\mu||\beta_n| \underset{\substack{(4.33)\\(4.34)}}{<} \frac{|\lambda| + |\mu|}{c}\varepsilon < \varepsilon.$$

Это и означает, что

$$\lim_{n \to \infty} (\lambda\alpha_n + \mu\beta_n) = 0,$$

т. е. что последовательность $\{\lambda\alpha_n + \mu\beta_n\}$ — бесконечно малая.

Соответствующее утверждение для любой конечной линейной комбинации бесконечно малых следует из доказанного методом математической индукции. □

Задача 4. Определив сумму бесконечного числа занумерованных слагаемых (обобщающую понятие суммы конечного числа слагаемых), а затем сумму бесконечного числа последовательностей, построить пример бесконечного числа бесконечно малых последовательностей, сумма которых не является бесконечно малой последовательностью.

II. *Произведение бесконечно малой последовательности на ограниченную последовательность является бесконечно малой последовательностью.*

Доказательство. Пусть $\{\alpha_n\}$ — бесконечно малая последовательность, а $\{x_n\}$ — ограниченная последовательность, т. е. существует такое число $b > 0$, что для всех номеров $n = 1, 2, \ldots$ выполняется неравенство $|x_n| \leqslant b$.

Зададим $\varepsilon > 0$; в силу определения бесконечно малой последовательности, существует такой номер n_ε, что для всех

$n > n_\varepsilon$ выполняется неравенство $|\alpha_n| < \frac{\varepsilon}{b}$. Поэтому для всех $n > n_\varepsilon$ имеем

$$|\alpha_n x_n| = |\alpha_n||x_n| < \frac{\varepsilon}{b} \cdot b = \varepsilon,$$

что и означает, что последовательность $\{\alpha_n x_n\}$ бесконечно малая. □

СЛЕДСТВИЕ. *Произведение конечного числа бесконечно малых последовательностей является бесконечно малой последовательностью.*

Это сразу следует по индукции из свойства II, если заметить, что бесконечно малая последовательность, как и всякая последовательность, имеющая предел, ограничена (см. теорему 2 п. 4.4).

Задача 5. Определив произведение бесконечного числа занумерованных сомножителей (обобщающее понятие произведения конечного числа сомножителей), а затем произведение бесконечного числа последовательностей, построить пример бесконечного числа бесконечно малых последовательностей, произведение которых не является бесконечно малой последовательностью.

УПРАЖНЕНИЕ 13. Доказать, что, для того чтобы последовательность $x_n \neq 0$, $n = 1, 2, \ldots$, была бесконечно малой, необходимо и достаточно, чтобы последовательность $\frac{1}{x_n}$, $n = 1, 2, \ldots$, была бесконечно большой.

4.9. Свойства пределов, связанные с арифметическими операциями над последовательностями

ЛЕММА. *Для того чтобы число a являлось пределом числовой последовательности $\{x_n\}$, необходимо и достаточно, чтобы ее члены x_n имели вид $x_n = a + \alpha_n$, $n = 1, 2, \ldots$, где $\{\alpha_n\}$ — бесконечно малая последовательность.*

В самом деле, пусть заданы какая-либо последовательность $\{x_n\}$ и число a; положим $\alpha_n \overset{\text{def}}{=} x_n - a$. Тогда условие $\lim\limits_{n \to \infty} x_n = a$, согласно определению предела последовательности, равносильно следующему: для любого $\varepsilon > 0$ существует такое $n_0 \in N$, что для всех $n > n_0$, $n \in N$, выполняется нера-

венство $|x_n - a| < \varepsilon$, т. е. неравенство $|\alpha_n| < \varepsilon$, а это и равносильно тому, что $\lim\limits_{n \to \infty} \alpha_n = 0$. □

Эта лемма показывает особую роль бесконечно малых последовательностей при изучении понятия предела, так как общее понятие предела последовательности с помощью этой леммы сводится к понятию нулевого предела. Это обстоятельство далее широко используется при изучении ряда свойств сходящихся последовательностей.

Докажем свойства пределов последовательностей, связанные с некоторыми операциями над членами числовых последовательностей.

1^0. *Если последовательность* $\{x_n\}$ *сходится, то сходится и последовательность* $\{|x_n|\}$, *причем если* $\lim\limits_{n \to \infty} x_n = a$, *то*

$$\lim_{n \to \infty} |x_n| = |a|.$$

Доказательство. Если $\lim\limits_{n \to \infty} x_n = a$, то для каждого $\varepsilon > 0$ существует такой номер n_ε, что для всех номеров $n > n_\varepsilon$ выполняется неравенство $|x_n - a| < \varepsilon$, но $||x_n| - |a|| < |x_n - a|$. Следовательно, для всех номеров $n > n_\varepsilon$ имеет место неравенство $||x_n| - |a|| < \varepsilon$, а это и означает, что $\lim\limits_{n \to \infty} |x_n| = |a|$. □

2^0. *Конечная линейная комбинация сходящихся последовательностей также является сходящейся последовательностью, и ее предел равен такой же линейной комбинации пределов данных последовательностей.*

Доказательство. Пусть

$$\lim_{n \to \infty} x_n = a \in \boldsymbol{R}, \quad \lim_{n \to \infty} y_n = b \in \boldsymbol{R}. \tag{4.35}$$

Тогда, в силу необходимости условий леммы для существования конечного предела, члены последовательностей $\{x_n\}$ и $\{y_n\}$ можно представить в виде

$$x_n = a + \alpha_n,\ y_n = b + \beta_n,\ n = 1, 2, \ldots, \tag{4.36}$$

где $\{\alpha_n\}$ и $\{\beta_n\}$ — бесконечно малые:

$$\lim_{n \to \infty} \alpha_n = \lim_{n \to \infty} \beta_n = 0. \tag{4.37}$$

Пусть теперь λ и μ — какие-либо числа. Тогда члены последовательности $\{\lambda x_n + \mu y_n\}$ представимы в виде

$$\lambda x_n + \mu y_n \underset{(4.36)}{=} (\lambda a + \mu b) + (\lambda\alpha_n + \mu\beta_n),\ n = 1, 2, \ldots, \quad (4.38)$$

где последовательность $\{\lambda\alpha_n + \mu\beta_n\}$, в силу бесконечной малости последовательностей $\{\alpha_n\}$ и $\{\beta_n\}$, также бесконечно малая (см. свойство 1^0 бесконечно малых последовательностей в п. 4.8):

$$\lim_{n \to \infty}(\lambda\alpha_n + \mu\beta_n) \underset{(4.37)}{=} 0. \quad (4.39)$$

Поэтому, в силу достаточности условий леммы для существования конечного предела, из равенств (4.38) следует, что последовательность $\{\lambda x_n + \mu y_n\}$ имеет предел, равный $\lambda a + \mu b$:

$$\lim_{n \to \infty}(\lambda x_n + \mu y_n) = \lambda a + \mu b,$$

т. е. (см. (4.35))

$$\lim_{n \to \infty}(\lambda x_n + \mu y_n) = \lambda \lim_{n \to \infty} x_n + \mu \lim_{n \to \infty} y_n .$$

Соответствующее утверждение для любой конечной линейной комбинации сходящихся последовательностей следует из доказанного, если воспользоваться методом математической индукции. □

3^0. *Если последовательности $\{x_n\}$ и $\{y_n\}$ сходятся, то их произведение $\{x_n y_n\}$ также сходится и*

$$\lim_{n \to \infty} x_n y_n = \lim_{n \to \infty} x_n \lim_{n \to \infty} y_n ,$$

т. е. предел произведения сходящихся последовательностей существует и равен произведению пределов данных последовательностей.

Доказательство. Пусть $\lim\limits_{n \to \infty} x_n = a$, $\lim\limits_{n \to \infty} y_n = b$; тогда

$$x_n = a + \alpha_n,\ y_n = b + \beta_n,\ n = 1, 2, \ldots,$$

где $\lim\limits_{n \to \infty} \alpha_n = \lim\limits_{n \to \infty} \beta_n = 0$; поэтому

$$x_n y_n = (a + \alpha_n)(b + \beta_n) = ab + (\alpha_n b + \beta_n a + \alpha_n\beta_n).$$

В силу свойств I и II бесконечно малых последовательностей (см. п. 4.8), $\lim\limits_{n \to \infty} (\alpha_n b + \beta_n a + \alpha_n \beta_n) = 0$; поэтому

$$\lim_{n \to \infty} x_n y_n = ab = \lim_{n \to \infty} x_n \lim_{n \to \infty} y_n . \square$$

СЛЕДСТВИЕ 1. *Если последовательность $\{x_n\}$ сходится, то для любого числа c последовательность $\{cx_n\}$ также сходится и*

$$\lim_{n \to \infty} cx_n = c \lim_{n \to \infty} x_n ,$$

т. е. постоянную можно выносить за знак предела.

Это утверждение сразу вытекает из свойства 3^0, когда последовательность $\{y_n\}$ стационарная: $y_n = c$, $n = 1, 2, \ldots$ (а также и из свойства 2^0, когда $y_n = 0$, $n = 1, 2, \ldots$).

СЛЕДСТВИЕ 2. *Если $\{x_n\}$ — сходящаяся последовательность и k — натуральное число, то*

$$\lim_{n \to \infty} (x_n)^k = (\lim_{n \to \infty} x_n)^k .$$

Это утверждение следует по индукции из свойства 3^0.

4^0. *Если последовательности $\{x_n\}$ и $\{y_n\}$ сходятся, $y_n \neq 0$, $n = 1, 2, \ldots$, и $\lim\limits_{n \to \infty} y_n \neq 0$, то последовательность $\left\{\dfrac{x_n}{y_n}\right\}$ сходится и*

$$\lim_{n \to \infty} \frac{x_n}{y_n} = \frac{\lim\limits_{n \to \infty} x_n}{\lim\limits_{n \to \infty} y_n},$$

т. е. при сделанных предположениях предел частного сходящихся последовательностей существует и равен частному от пределов данных последовательностей.

Доказательство. Пусть $\lim\limits_{n \to \infty} x_n = a$, $\lim\limits_{n \to \infty} y_n = b \neq 0$ и для определенности $b > 0$. Тогда

$$x_n = a + \alpha_n, \quad y_n = b + \beta_n, \quad n = 1, 2, \ldots ,$$

где $\lim\limits_{n \to \infty} \alpha_n = \lim\limits_{n \to \infty} \beta_n = 0$, а согласно следствию 1 из свойства III пределов последовательностей в п. 4.3, существует та-

кой номер n_0, что для всех номеров $n > n_0$ выполняется неравенство $y_n > \frac{b}{2} > 0$ (действительно, $\frac{b}{2} < b$, здесь используется предположение, что $b > 0$); поэтому при $n > n_0$ имеем $\frac{1}{y_n} < \frac{2}{b}$ ($y_n \neq 0$, поэтому на него можно делить).

Далее

$$\frac{x_n}{y_n} - \frac{a}{b} = \frac{a + \alpha_n}{b + \beta_n} - \frac{a}{b} = \frac{1}{b(b + \beta_n)}(\alpha_n b - \beta_n a). \qquad (4.40)$$

Здесь $0 < \frac{1}{b(b + \beta_n)} = \frac{1}{by_n} < \frac{2}{b^2}$, т. е. последовательность $\frac{1}{b(b + \beta_n)}$, $n = n_0 + 1, n_0 + 2, \ldots$, ограничена (отсюда, конечно, следует, что эта последовательность ограничена и при всех $n = 1, 2, \ldots$).

Согласно свойствам бесконечно малых последовательностей, последовательность $\{\alpha_n b - \beta_n a\}$ является бесконечно малой, поэтому и последовательность $\left\{\frac{1}{b(b + \beta_n)}(\alpha_n b - \beta_n a)\right\}$ бесконечно малая. В силу этого, из равенства (4.40) следует, что

$$\lim_{n \to \infty} \frac{x_n}{y_n} = \frac{a}{b} = \frac{\lim\limits_{n \to \infty} x_n}{\lim\limits_{n \to \infty} y_n}.$$

Аналогично рассматривается случай, когда $b < 0$. □

З а м е ч а н и е 1. В случае последовательностей, имеющих бесконечные пределы, утверждения, аналогичные свойствам 2^0—4^0, вообще говоря, не имеют места. Например, пусть $x_n = n + 1$, $y_n = n$, $n = 1, 2, \ldots$; тогда

$$\lim_{n \to \infty} x_n = \lim_{n \to \infty} y_n = +\infty \quad \text{и} \quad \lim_{n \to \infty} (x_n - y_n) = 1.$$

Если $x_n = 2n$, $y_n = n$, $n = 1, 2, \ldots$, то

$$\lim_{n \to \infty} x_n = \lim_{n \to \infty} y_n = +\infty \quad \text{и} \quad \lim_{n \to \infty} (x_n - y_n) = +\infty.$$

Если же $x_n = n + \sin\frac{n\pi}{2}$, $y_n = n$, $n = 1, 2, \ldots$, то

$$\lim_{n \to \infty} x_n = \lim_{n \to \infty} y_n = +\infty,$$

а последовательность $x_n - y_n = \sin\frac{\pi n}{2}$, $n = 1, 2, \ldots$, не имеет ни конечного, ни бесконечного предела.

Эти примеры показывают, что при одинаковых предположениях о последовательностях $\{x_n\}$ и $\{y_n\}$, имеющих бесконечные пределы, для последовательностей $\{x_n - y_n\}$ могут встретиться самые разнообразные случаи. Вместе с тем отдельные обобщения свойств 2^0—4^0 на случай последовательностей с бесконечными пределами все-таки имеют место. Например, если $\lim\limits_{n\to\infty} x_n = +\infty$ и $\lim\limits_{n\to\infty} y_n = +\infty$ (или $\lim\limits_{n\to\infty} y_n$ конечен), то $\lim\limits_{n\to\infty}(x_n + y_n) = +\infty$, или если $\alpha > 0$ и $\lim\limits_{n\to\infty} x_n = +\infty$, то $\lim\limits_{n\to\infty} \alpha x_n = \alpha \lim\limits_{n\to\infty} x_n = +\infty$ (рекомендуется доказать самостоятельно), что оправдывает операции, введенные для бесконечностей $+\infty$ и $-\infty$ в п. 3.1.

З а м е ч а н и е 2. Если разность $\{x_n - y_n\}$ последовательностей $\{x_n\}$ и $\{y_n\}$ является бесконечно малой, то они одновременно имеют или не имеют конечные пределы, причем если эти пределы существуют, то они равны.

В самом деле, пусть $x_n - y_n = \alpha_n$, $n = 1, 2, \ldots$, где $\lim\limits_{n\to\infty} \alpha_n = 0$. Если существует конечный предел $\lim\limits_{n\to\infty} x_n = a$, то существует и конечный предел

$$\lim_{n\to\infty} y_n = \lim_{n\to\infty}(x_n - \alpha_n) = \lim_{n\to\infty} x_n - \lim_{n\to\infty} \alpha_n = a.$$

Последовательности $\{x_n\}$ и $\{y_n\}$ равноправны, поэтому утверждение доказано.

УПРАЖНЕНИЯ. 14. Доказать, что если разность $\{x_n - y_n\}$ последовательностей $\{x_n\}$ и $\{y_n\}$ является бесконечно малой, то они одновременно имеют или не имеют бесконечные пределы, причем если эти пределы существуют, то они равны.

15. Доказать, что если $\lim\limits_{n\to\infty} x_n = +\infty$, а последовательность $\{y_n\}$ ограничена, то

$$\lim_{n\to\infty}(x_n + y_n) = +\infty.$$

Пример 1. Пусть $a > 0$, $x_0 > 0$ и

$$x_n \stackrel{\text{def}}{=} \frac{1}{2}\left(x_{n-1} + \frac{a}{x_{n-1}}\right),\ n = 1, 2, \ldots . \tag{4.41}$$

Докажем, что $\lim\limits_{n \to \infty} x_n = \sqrt{a}$. По индукции сразу ясно, что $x_n > 0$ для всех $n = 0, 1, 2, \ldots$. Более того, покажем, что

$$x_n \geqslant \sqrt{a},\ n = 1, 2, \ldots . \tag{4.42}$$

Для этого предварительно заметим, что из очевидного неравенства $(t-1)^2 \geqslant 0$ в случае $t > 0$ следует неравенство $t + \frac{1}{t} \geqslant 2$. Используя это неравенство при $t = \frac{x_n}{\sqrt{a}}$, в силу (4.41) получаем

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) = \frac{\sqrt{a}}{2}\left(\frac{x_n}{\sqrt{a}} + \frac{\sqrt{a}}{x_n}\right) \geqslant \frac{\sqrt{a}}{2}2 = \sqrt{a},$$

$$n = 0, 1, 2, \ldots .$$

Покажем теперь, что последовательность x_n, $n = 1, 2, \ldots$, убывает. Применяя неравенство (4.42), получаем

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) =$$

$$= \frac{x_n}{2}\left(1 + \frac{a}{x_n^2}\right) \leqslant \frac{x_n}{2}2 = x_n, \quad n = 1, 2, \ldots . \tag{4.43}$$

Итак, $\sqrt{a} \leqslant \ldots \leqslant x_{n+1} \leqslant x_n \leqslant \ldots \leqslant x_1$, где бы ни было расположено «нулевое приближение» $x_0 > 0$, т. е. последовательность $\{x_n\}$ ограничена снизу и монотонно убывает, поэтому, согласно теореме 3, она имеет предел.

Пусть $\lim\limits_{n \to \infty} x_n = x$. Переходя к пределу в равенстве (4.41) при $n \to \infty$, получаем равенство $x = \frac{1}{2}\left(x + \frac{a}{x}\right)$, откуда $x^2 = a$, и так как $x_n \geqslant 0$, то и $x \geqslant 0$, поэтому $x = \sqrt{a}$.

Формула (4.41) может служить для приближенного вычисления значений квадратного корня из числа a. Ее действительно применяют с этой целью на практике.

Нетрудно подсчитать и точность, с которой n-е приближение, т. е. член x_n, дает значение корня $\sqrt{a}$.

Из рекуррентной формулы (4.41) имеем

$$x_{n+1} - \sqrt{a} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) - \sqrt{a} =$$

$$= \frac{1}{2x_n}(x_n^2 - 2x_n\sqrt{a} + a) = \frac{1}{2x_n}(x_n - \sqrt{a})^2,\ n = 0, 1, 2, \ldots .$$

Применяя неравенство (4.42), отсюда находим

$$0 \leqslant x_{n+1} - \sqrt{a} \leqslant \frac{1}{2\sqrt{a}}(x_n - \sqrt{a})^2,\ n = 1, 2, \ldots .$$

Полученная оценка не совсем удобна на практике, поскольку мы не знаем значения корня $\sqrt{a}$ — мы его ищем. Однако всегда можно найти приближенно такое c, что $0 < c < \sqrt{a}$, причем можно выбрать и $x_0 \geqslant c$, тогда из полученной оценки будем иметь

$$0 \leqslant x_{n+1} - \sqrt{a} \leqslant \frac{1}{2c}(x_n - \sqrt{a})^2,\ \ n = 0, 1, 2, \ldots,$$

или

$$\frac{1}{2c}(x_{n+1} - \sqrt{a}) \leqslant \left[\frac{1}{2c}(x_n - \sqrt{a})\right]^2,\ \ n = 0, 1, 2, \ldots .$$

Отсюда по индукции находим

$$\frac{1}{2c}(x_n - \sqrt{a}) \leqslant \left[\frac{1}{2c}(x_{n-1} - \sqrt{a})\right]^2 \leqslant$$

$$\leqslant \left[\left(\frac{1}{2c}(x_{n-2} - \sqrt{a})\right)^2\right]^2 \leqslant \ldots \leqslant \left[\frac{1}{2c}(x_0 - \sqrt{a})\right]^{2^n}. \qquad (4.44)$$

Если выбрать нулевое приближение x_0 так, чтобы

$$q \overset{\text{def}}{=} \frac{1}{2c}|x_0 - \sqrt{a}| < 1,$$

то из (4.44) получится, что

$$0 \leqslant x_n - \sqrt{a} \leqslant 2cq^{2^n},\ \ n = 1, 2, \ldots ,$$

т. е. последовательность (4.41) сходится к значению корня гораздо быстрее геометрической прогрессии со знаменателем $0 < q < 1$.

Для иллюстрации приближенного вычисления корня по формуле (4.41) приведем результаты вычисления $\sqrt{2}$ на ЭВМ в случае, когда в качестве нулевого приближения x_0 выбрано $x_0 = 1$:

x_1 = **1,50000000000000000000000000000000000000000000000000000000000**,

x_2 = **1,41**666666666666666666666666666666666666666666666666666666666,

x_3 = **1,41421**568627450980392156862745098039215686274509803921568627,

x_4 = **1,41421356237**468991062629557889013491011655962211574404458490,

x_5 = **1,41421356237309504880168**962350253024361498192577619742849828,

x_6 = **1,414213562373095048801688724209698078569671875**37723400156101,

x_7 = **1,41421356237309504880168872420969807856967187537694807317667**.

Здесь полужирным шрифтом выделены числа, являющиеся правильным приближением $\sqrt{2}$ с точностью до числа значащих цифр, входящих в выделенное число (эти цифры стабилизируются в процессе вычисления). Число правильных в этом смысле цифр удваивается (как это видно из таблицы) при переходе к следующему приближению: в x_1 имеется одна правильная цифра, в x_2 — уже 3, в x_3 — 6, в x_4 — 12, в x_5 — 24, в x_6 — 48, а x_7 дает уже правильное приближение $\sqrt{2}$ с 60 значащими цифрами, т. е. со всеми использованными разрядами.

Заметим, что если число $a > 0$ таково, что корень из него вычисляется в конечном виде в том смысле, что существует такая конечная десятичная дробь x, что $x^2 = a$, то эта дробь может быть найдена (во всяком случае, принципиально) с помощью классического метода извлечения корня «столбиком». В случае же извлечения корня из этого числа с помощью формулы (4.41), если x_0 выбрано отличным от точного значения корня: $x_0 \neq \sqrt{a}$, то указанное точное значение корня не получится ни на каком шаге. Это следует из того, что в случае $x_0 \neq \sqrt{a}$ последовательность (4.41) строго убывает: $x_{n+1} < x_n$, $n = 1, 2, \ldots$.

Прежде чем рассматривать дальнейшие примеры, докажем следующее полезное неравенство.

ЛЕММА (неравенство Бернулли[1]). *Пусть* $\alpha > 0$, *тогда для любого натурального n справедливо неравенство*

$$(1 + \alpha)^n > 1 + n\alpha. \qquad (4.45)$$

Доказательство. По формуле бинома Ньютона имеем

$$(1 + \alpha)^n = 1 + n\alpha + \frac{n(n-1)}{2}\alpha^2 + \ldots + \alpha^n.$$

Все слагаемые в правой части равенства положительны: отбросим все из них, кроме первого и второго, в результате правая часть может только уменьшиться и мы получим исходное неравенство (4.45). □

Вернемся к рассмотрению примеров.

Пример 2. Если $p > 1$, то $\lim\limits_{n \to \infty} p^n = +\infty$, а если $0 < p < 1$, то $\lim\limits_{n \to \infty} p^n = 0$.

Пусть сначала $p > 1$; тогда $p = 1 + \alpha$, где $\alpha > 0$, и, согласно неравенству Бернулли (см. (4.19)),

$$p^n = (1 + \alpha)^n > 1 + n\alpha > n\alpha.$$

Так как $\lim\limits_{n \to \infty} \alpha n = \alpha \lim\limits_{n \to \infty} n = +\infty$, $\alpha > 0$, то и $\lim\limits_{n \to \infty} p^n = +\infty$. Если теперь $0 < p < 1$, то $q = \frac{1}{p} > 1$ и

$$\lim_{n \to \infty} p^n = \lim_{n \to \infty} \frac{1}{q^n} = \frac{1}{\lim\limits_{n \to \infty} q^n} = 0,$$

ибо, по доказанному, $\lim\limits_{n \to \infty} q^n = +\infty$.

Пример 3. Для любого $a > 0$

$$\lim_{n \to \infty} a^{1/n} = 1, \qquad (4.46)$$

$$\lim_{n \to \infty} a^{-1/n} = 1. \qquad (4.47)$$

---

[1] Я. Бернулли (1654—1705) — швейцарский математик.

Пусть сначала $a > 1$, тогда $b \stackrel{\text{def}}{=} \sqrt[n]{a} > 1$. В самом деле, согласно определению корня, $b^n = a$. Если $b \leqslant 1$, то, перемножив это неравенство n раз, мы получим, что $a = b^n \leqslant 1$, но это противоречит условию $a > 1$. Положим

$$x_n \stackrel{\text{def}}{=} \sqrt[n]{a} - 1. \tag{4.48}$$

Согласно сказанному, $x_n > 0$. Из (4.48) следует, что $a = (1 + x_n)^n$. Применив неравенство Бернулли, получим

$$a = (1 + x_n)^n > 1 + nx_n > nx_n.$$

Следовательно, $0 < x_n < \frac{a}{n}$, поэтому $\lim\limits_{n \to \infty} x_n = 0$; откуда, согласно (4.48),

$$\lim_{n \to \infty} \sqrt[n]{a} = 1.$$

Если теперь $0 < a < 1$, то $b \stackrel{\text{def}}{=} \frac{1}{a} > 1$, и так как, в силу доказанного, $\lim\limits_{n \to \infty} \sqrt[n]{b} = 1$, то

$$\lim_{n \to \infty} \sqrt[n]{a} = \lim_{n \to \infty} \sqrt[n]{\frac{1}{b}} = \lim_{n \to \infty} \frac{1}{\sqrt[n]{b}} = \frac{1}{\lim\limits_{n \to \infty} \sqrt[n]{b}} = 1.$$

Если $a = 1$, то $\sqrt[n]{a} = 1$, $n = 1, 2, \ldots$, и, следовательно, также $\lim\limits_{n \to \infty} \sqrt[n]{a} = 1$.

Таким образом, (4.46) доказано при любом $a > 0$. Отсюда сразу следует (4.47):

$$\lim_{n \to \infty} a^{-1/n} = \lim_{n \to \infty} \frac{1}{a^{1/n}} = \frac{1}{\lim\limits_{n \to \infty} a^{1/n}} = 1. \ \square$$

Пример 4. $$\lim_{n \to \infty} \frac{a^n}{n!} = 0, \ a > 0. \tag{4.49}$$

Действительно, пусть $n_0 \in N$ таково, что $\frac{a}{n_0} < \frac{1}{2}$. Тогда при всех $n > n_0$ справедливо неравенство $\frac{a}{n} < \frac{1}{2}$, поэтому

$$\frac{a^n}{n!} = \frac{a^{n_0}}{n_0!} \frac{a}{n_0 + 1} \frac{a}{n_0 + 2} \ldots \frac{a}{n} < \frac{a^{n_0}}{n_0!} \left(\frac{1}{2}\right)^{n - n_0}.$$

Заметив, что $\lim\limits_{n\to\infty}\frac{a^{n_0}}{n_0!}\left(\frac{1}{2}\right)^{n-n_0} = \frac{2^{n_0}a^{n_0}}{n_0!}\lim\limits_{n\to\infty}\left(\frac{1}{2}\right)^n = 0$, получим

$$\lim_{n\to\infty}\frac{a^n}{n!} = 0.$$

Равенство (4.49) можно доказать и иначе. Рассмотрим последовательность $x_n = \frac{a^n}{n!}$; тогда $x_{n+1} = x_n\frac{a}{n+1}$ и при $n + 1 > a$ выполняется неравенство $x_{n+1} < x_n$, т. е. начиная с некоторого номера последовательность $\{x_n\}$ убывает. Кроме того, при всех $n \in N$ имеет место неравенство $x_n > 0$, поэтому эта последовательность ограничена снизу и, следовательно, имеет конечный предел. Пусть $x = \lim\limits_{n\to\infty} x_n$. Переходя к пределу при $n \to \infty$ в равенстве $x_{n+1} = x_n\frac{a}{n+1}$, получим

$$x = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} x_n \lim_{n\to\infty}\frac{a}{n+1} = x\cdot 0 = 0,$$

т. е. снова имеем (4.49).

Пример 5. $$\lim_{n\to\infty}\sqrt[n]{n!} = +\infty. \qquad (4.50)$$

При любом $a > 0$ имеет место равенство (4.49), поэтому для любого $a > 0$ существует такое $n_a \in N$, что для всех $n > n_a$ выполняется неравенство

$$\frac{a^n}{n!} < 1,$$

т. е. при $n > n_a$ имеет место $\sqrt[n]{n!} > a$, и так как $a > 0$ произвольно, то это и означает справедливость равенства (4.50).

Пример 6. $$\lim_{n\to\infty}\left(1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!}\right) = e. \qquad (4.51)$$

Напомним (см. п. 4.5), что число e является пределом последовательности

$$x_n = \left(1 + \frac{1}{n}\right)^n,\ n = 1, 2, \ldots,$$

т. е.

$$\lim_{n\to\infty} x_n = e, \qquad (4.52)$$

причем в силу строгого возрастания последовательности $\{x_n\}$ имеет место неравенство

$$x_n < e. \qquad (4.53)$$

Положим

$$s_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!}.$$

Было показано (см. (4.28)), что

$$x_n < s_n, \ \ n = 1, 2, \ldots . \qquad (4.54)$$

С другой стороны, зафиксировав в формуле (4.27) произвольное $k \geqslant 1$ и выбрав $n > k$, отбросим в правой части равенства (4.27) все слагаемые начиная с $(k + 2)$-го. В результате получим неравенство

$$x_n > 1 + \frac{1}{1!} + \frac{1}{2!}\left(1 - \frac{1}{n}\right) + \ldots + \frac{1}{k!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ldots\left(1 - \frac{k-1}{n}\right).$$

Перейдя в этом неравенстве к пределу при $n \to \infty$ и фиксированном k и заметив, что правая часть имеет своим пределом s_k, получим, в силу (4.52), неравенство

$$e \geqslant s_k, \ \ k = 1, 2, \ldots . \qquad (4.55)$$

Объединив (4.54) и (4.55), получим

$$x_n < s_n \leqslant e, \ \ n = 1, 2, \ldots .$$

Отсюда, согласно (4.52), непосредственно следует, что $\lim\limits_{n \to \infty} s_n = e$, т. е. равенство (4.51).

З а м е ч а н и е. Для приближенного вычисления числа e формула $e \approx \left(1 + \frac{1}{n}\right)^n$ не очень удобна, так как при переходе от n к $n + 1$ приходится все вычисления производить заново. Приближенная формула $e \approx s_n$ более удобна для числовых расчетов, ибо при переходе от n к $n + 1$ надо к уже найденному значению s_n прибавить число $\frac{1}{(n+1)!}$:

$$s_{n+1} = s_n + \frac{1}{(n+1)!},$$

т. е. проделанные при нахождении значения s_n вычисления не пропадают зря. Кроме того, в этом случае легко установить и оценку погрешности при замене числа e на значение суммы s_n. (Это будет сделано в п. 34.14*, пример 4.)

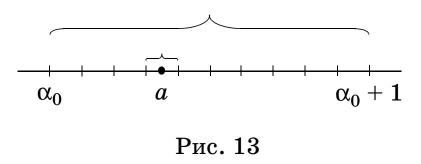


Рис. 13

УПРАЖНЕНИЕ 16. Положим $a_0 > 0$, $b_0 \geqslant 0$, $a_n = \sqrt{a_{n-1}b_{n-1}}$, $b_n = \frac{a_{n-1} + b_{n-1}}{2}$, $n = 1, 2, \ldots$. Доказать, что последовательности $\{a_n\}$ и $\{b_n\}$ стремятся к одному и тому же пределу a и что $0 \leqslant a - a_n \leqslant \frac{|b_0 - a_0|}{2^n}$, $0 \leqslant b_n - a \leqslant \frac{|b_0 - a_0|}{2^n}$.

4.10. Изображение действительных чисел бесконечными десятичными дробями

Пусть задано какое-либо число a, для определенности $a \geqslant 0$. Согласно принципу Архимеда, существует целое число $n_0 > a$. Среди чисел $n = 1, 2, \ldots, n_0$ возьмем наименьшее, обладающее свойством $n > a$, и обозначим его $\alpha_0 + 1$, тогда $\alpha_0 \leqslant a < \alpha_0 + 1$.

Разобьем отрезок $I_0 = [\alpha_0; \alpha_0 + 1]$ на десять равных отрезков, т. е. рассмотрим отрезки $\left[\alpha_0,\ \alpha_1;\ \alpha_0,\ \alpha_1 + \frac{1}{10}\right]$, где α_0, α_1 — десятичная дробь, причем α_1 обозначает порядковый номер отрезка, содержащего число a и получившегося при разбиении отрезка I_0 на десять равных отрезков при последовательной их нумерации слева направо числами $0, 1, 2, \ldots, 9$. В качестве разделительного знака между концами отрезков будем употрелять здесь не запятую, как обычно, а точку с запятой, чтобы отличать этот знак от запятой в десятичной дроби.

Возможны два случая: либо точка a не совпадает ни с одной точкой деления (рис. 13), либо точка a совпадает с одной

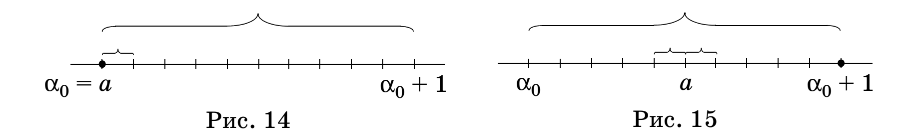


Рис. 14 Рис. 15

из точек деления (рис. 14, 15). В первом случае точка a принадлежит только одному из этих отрезков. Обозначим его I_1, т. е. $I_1 = \left[\alpha_0, \alpha_1; \alpha_0, \alpha_1 + \frac{1}{10}\right]$. Во втором случае точка a может принадлежать двум соседним отрезкам (рис. 15). Тогда через $I_1 = \left[\alpha_0, \alpha_1; \alpha_0, \alpha_1 + \frac{1}{10}\right]$ обозначим тот из них, для которого точка a является левым концом. В обоих случаях $a \in I_1$. Разобьем отрезок I_1, в свою очередь, на десять равных отрезков и через $I_2 = \left[\alpha_0, \alpha_1\alpha_2; \alpha_0, \alpha_1\alpha_2 + \frac{1}{10^2}\right]$ обозначим тот из получившихся отрезков, который содержит a и для которого точка a не является правым концом. Продолжив этот процесс, получим систему вложенных отрезков

$$I_n = [\underline{a}_n, \overline{a}_n], \ \ n = 0, 1, 2, \ldots ,$$

где

$$\underline{a}_n = \alpha_0, \alpha_1\alpha_2\ldots , \ \ \overline{a}_n = \alpha_0, \alpha_1\alpha_2\ldots\alpha_n + \frac{1}{10^n},$$

а α_n — одна из цифр 0, 1, 2, ... , 9. Каждый из отрезков I_n содержит a, причем a не является его правым концом,

$$a \in I_n, \ \ a \neq \overline{a}_n, \ \ n = 0, 1, 2, \ldots;$$

длина отрезка I_n равна 10^{-n} и, следовательно, стремится к нулю при $n \to \infty$.

Конечные десятичные дроби $\underline{a}_n$ и $\overline{a}_n$ называются *десятичными дробями, приближающими число* a. Более точно, число $\underline{a}_n$ *называется нижним десятичным приближением* порядка n, а число $\overline{a}_n$ — *верхним десятичным приближением* того же порядка числа a. Они обладают следующими свойствами, непосредственно вытекающими из их определения:

$$\underline{a}_n \leqslant a < \overline{a}_n, \tag{4.56}$$

$$\underline{a}_n \leqslant \underline{a}_{n+1}, \overline{a}_{n+1} \leqslant \overline{a}_n, \tag{4.57}$$

$$\overline{a}_n - \underline{a}_n = 10^{-n}. \tag{4.58}$$

В том случае, если $a < 0$, полагая $b = -a$, определяем

$$\underline{a}_n = -\overline{b}_n, \ \overline{a}_n = -\underline{b}_n,$$

при этом свойства (4.56)—(4.58), очевидно, сохраняются, лишь в неравенстве (4.56) знаки $\leqslant$ и $<$ поменяются местами.

Свойство (4.57) означает, что отрезки $[\underline{a}_n, \overline{a}_n]$ образуют вложенную систему отрезков. Из свойства (4.58) следует, что длины отрезков $[\underline{a}_n, \overline{a}_n]$ стремятся к нулю. Наконец, (4.56) означает, что точка a принадлежит всем этим отрезкам, поэтому, согласно замечанию 2 п. 4.5, она является пределом их концов $\underline{a}_n$ и $\overline{a}_n$.

Итак, в частности, доказана следующая лемма.

ЛЕММА 1. *Каково бы ни было число a, последовательность $\{\underline{a}_n\}$ возрастает, последовательность $\{\overline{a}_n\}$ убывает и*

$$\lim_{n \to \infty} \underline{a}_n = \lim_{n \to \infty} \overline{a}_n = a.$$

СЛЕДСТВИЕ. *Всякое действительное число является пределом последовательности рациональных чисел.*

Следствие леммы вытекает из того, что $\underline{a}_n$ и $\overline{a}_n$ суть рациональные числа.

Пусть теперь снова $a \geqslant 0$ и $\underline{a}_n = \alpha_0, \alpha_1\alpha_2...\alpha_n$. Поставим в соответствие числу a бесконечную десятичную дробь $\alpha_0, \alpha_1\alpha_2...\alpha_n...$. Подчеркнем, что здесь α_0 является неотрицательным целым числом, а α_n, $n = 1, 2, \ldots$, — одной из цифр 0, 1, 2, ..., 9. Число a является единственным числом, принадлежащим всем отрезкам I_n, $n = 1, 2, \ldots$, поэтому при указанном соответствии разным числам соответствуют разные десятичные дроби, т. е. отличающиеся хотя бы одним α_k ($k = 0, 1, 2, \ldots$).

Заметим далее, что при таком построении не может получиться дробь с периодом, состоящим из одной цифры 9. Действительно, пусть числу a соответствует следующая дробь: $\alpha_0, \alpha_1...\alpha_{n_0}\ 9...9...$, где в случае $n_0 \neq 0$ выполняется неравенство $\alpha_{n_0} \neq 9$. Тогда, согласно построению,

$$a \in \left[\alpha_0, \alpha_1 \ldots \alpha_{n_0} 9 \ldots 9;\ \ \alpha_0, \alpha_1 \ldots \alpha_{n_0}\right]$$

для всех $n > n_0$, где n — число десятичных знаков после запятой в дроби $\alpha_0, \alpha_1, \dots \alpha_{n_0}$ 9 … 9. Отсюда следует, что a является правым концом всех отрезков I_n, $n > n_0$, что противоречит выбору этих отрезков.

Таким образом, в силу установленного соответствия каждому действительному числу $a \geqslant 0$ соответствует некоторая бесконечная десятичная дробь, не имеющая периода, состоящего из одной цифры 9. Такие десятичные дроби называются *допустимыми*.

Наконец, каждая бесконечная допустимая десятичная дробь $\alpha_0, \alpha_1, \alpha_2 \dots \alpha_n \dots$ в результате описанного соответствия оказывается поставленной в соответствие некоторому числу a, а именно тому единственному числу, которое принадлежит всем отрезкам:

$$\left[\alpha_0, \alpha_1 \dots \alpha_n;\ \ \alpha_0, \alpha_1 \dots \alpha_n + \frac{1}{10^n}\right],\quad n = 1, 2, \dots .$$

Это соответствие можно распространить и на отрицательные числа: если числу $a > 0$ соответствует дробь $\alpha_0, \alpha_1 \dots \alpha_n \dots$, то числу $-a$ поставим в соответствие дробь $-\alpha_0, \alpha_1 \dots \alpha_n \dots$, которую также будем называть допустимой.

Полученные результаты можно сформулировать в виде следующей теоремы.

ТЕОРЕМА 6. *Между множеством всех действительных чисел и множеством допустимых десятичных дробей существует взаимно однозначное соответствие; причем если при этом соответствии числу a соответствует дробь $\pm\alpha_0, \alpha_1\alpha_2 \dots \alpha_n \dots$, то*

$$\lim_{n \to \infty} \pm\alpha_0, \alpha_1\alpha_2 \dots \alpha_n = a.$$

Бесконечная десятичная дробь $\pm\alpha_0, \alpha_1\alpha_2 \dots \alpha_n \dots$, соответствующая числу a, называется его *десятичной записью* и используется для его обозначения; поэтому пишут

$$a = \pm\alpha_0, \alpha_1\alpha_2 \dots \alpha_n \dots .$$

Если бесконечная десятичная дробь имеет период, состоящий только из нулей: $\alpha_0, \alpha_1\alpha_2 \dots \alpha_n 0\ 0 \dots 0 \dots$, причем $\alpha_n \neq 0$, то говорят, что эта дробь имеет n значащих цифр после запятой; при этом обычно нуль в периоде не пишется,

т. е. указанное число записывается конечной десятичной дробью $\alpha_0, \alpha_1\alpha_2...\alpha_n$ (именно такая запись и употреблялась выше).

Замечание 1. Любой бесконечной десятичной дроби $\alpha_0, \alpha_1\alpha_2 \ldots \alpha_n \ldots$ (не обязательно допустимой) можно также естественным образом поставить в соответствие единственное действительное число, принадлежащее всем отрезкам:

$$\left[\alpha_0, \alpha_1, \ldots, \alpha_n;\ \alpha_0, \alpha_1, \ldots, \alpha_n + \frac{1}{10^n}\right].$$

Однако получившееся при этом соответствие уже не является взаимно однозначным: имеются разные бесконечные десятичные дроби, которым соответствует одно и то же действительное число. Именно дробям вида

$$\alpha_0, \alpha_1\alpha_2\ldots \alpha_n\, 99\ldots9\ldots \text{ и } \alpha_0, \alpha_1\alpha_2\ldots (\alpha_n + 1)\, 0\, 0\ldots0\ldots\ (\alpha_n \neq 9)$$

соответствует одно и то же число. В описанной выше конструкции соответствия действительных чисел и бесконечных десятичных дробей мы получили бы не только допустимые десятичные дроби, если бы отказались от условия каждый раз выбирать такой отрезок I_n, что число a не является его правым концом.

Используя запись действительных чисел, с помощью бесконечных десятичных дробей можно получить правило для их сравнения по величине и правила арифметических действий над ними. И то и другое сводится к аналогичным операциям над соответствующими их десятичными приближениями и, быть может, предельному переходу. Сформулируем эти результаты в виде лемм.

ЛЕММА 2. *Пусть a и b — действительные числа. Тогда $a < b$ в том и только в том случае, когда существует такое натуральное n_0, что для всех $n > n_0$ имеет место неравенство*

$$\underline{a}_n < \underline{b}_n.$$

Доказательство. Пусть $a < b$. Из $\lim\limits_{n\to\infty} \underline{a}_n = a$, $\lim\limits_{n\to\infty} \underline{b}_n = b$, согласно свойству III пределов последовательностей (см. п. 4.3), сразу следует существование требуемого номера n_0,

т. е. такого, что для всех номеров $n > n_0$ выполняется неравенство $\underline{a}_n < \underline{b}_n$.

Обратно: если указанный номер n_0 существует, то случай $a > b$ невозможен в силу только что доказанного. Невозможен и случай $a = b$, так как тогда бы, в силу однозначности записи чисел с помощью допустимых десятичных дробей, для всех $n = 1, 2, \ldots$ имело бы место равенство $\underline{a}_n = \underline{b}_n$. Следовательно, $a < b$. □

ЛЕММА 3. *Пусть a и b — два действительных числа, тогда*

$$\lim_{n \to \infty} (\underline{a}_n + \underline{b}_n) = a + b, \qquad \lim_{n \to \infty} (\underline{a}_n - \underline{b}_n) = a - b,$$

$$1 \lim_{n \to \infty} \underline{a}_n \underline{b}_n = ab,$$

а при $b \neq 0$

$$\lim_{n \to \infty} \frac{\underline{a}_n}{\underline{b}_n} = \frac{a}{b}.$$

Все утверждения этой леммы непосредственно следуют из леммы 1 и свойств пределов, связанных с арифметическими действиями над последовательностями (см. п. 4.9).

З а м е ч а н и е 2. Из леммы 3 следует, что, для того чтобы произвести с заданной степенью точности какое-либо арифметическое действие над числами, записанными в виде допустимых десятичных дробей, надо взять с достаточной точностью конечные десятичные приближения и выполнить соответствующие действия. При этом при сложении, вычитании и умножении в результате получается снова конечная десятичная дробь. В случае же деления частное двух конечных десятичных дробей будет, вообще говоря, бесконечной десятичной дробью, причем, как это известно из элементарной математики, периодической. Однако и в этом случае можно с любой степенью точности получить результат, выраженный конечной десятичной дробью. Например, если $\underline{(\underline{a}_n/\underline{b}_n)_n}$ является нижним десятичным приближением порядка n для частного $\left(\frac{\underline{a}_n}{\underline{b}_n}\right)$, то

$$\lim_{n \to \infty} \underline{\left(\frac{\underline{a}_n}{\underline{b}_n}\right)}_n = \frac{a}{b} \tag{4.59}$$

и, следовательно, частное a/b, $b \neq 0$, можно с любой степенью точности выразить с помощью конечных десятичных дробей вида $\left(\frac{\underline{a}_n}{\underline{b}_n}\right)_n$.

Для доказательства равенства (4.59) положим

$$\alpha_n = \frac{\underline{a}_n}{\underline{b}_n} - \frac{a}{b};$$

в силу леммы 3, имеем (см. лемму в п. 4.9) $\lim\limits_{n \to \infty} \alpha_n = 0$. Теперь, используя (4.56) и (4.58), получаем

$$\left(\frac{\underline{a}_n}{\underline{b}_n}\right)_n - \frac{a}{b} = \left[\left(\frac{\underline{a}_n}{\underline{b}_n}\right)_n - \frac{\underline{a}_n}{\underline{b}_n}\right] + \left(\frac{\underline{a}_n}{\underline{b}_n} - \frac{a}{b}\right) < \frac{1}{10^n} + \alpha_n.$$

И, поскольку $\lim\limits_{n \to \infty}\left(\frac{1}{10^n} + \alpha_n\right) = 0$, равенство (4.59) доказано. □

З а м е ч а н и е 3. В результате указанных выше вычислений с нижними десятичными приближениями порядка n в случае сложения $\underline{a}_n + \underline{b}_n$, вычитания $\underline{a}_n - \underline{b}_n$ и деления $\underline{(\underline{a}_n/\underline{b}_n)}_n$ мы снова получим конечные десятичные дроби с не более чем n значащими цифрами после запятой. При умножении же $\underline{a}_n \underline{b}_n$ получается, вообще говоря, десятичная дробь с $2n$ значащими цифрами после запятой. Если $\underline{(\underline{a}_n \underline{b}_n)}_n$ является нижним десятичным приближением произведения $\underline{a}_n \underline{b}_n$, то аналогично (4.59) доказывается, что

$$\lim_{n \to \infty} \underline{(\underline{a}_n \underline{b}_n)}_n = ab.$$

Таким образом, при приближенных вычислениях сумм $a + b$, разностей $a - b$, произведений ab и частных a/b, $b \neq 0$, соответственно по формулам

$$\underline{a}_n + \underline{b}_n,\ \underline{a}_n - \underline{b}_n,\ \underline{(\underline{a}_n \underline{b}_n)}_n \text{ и } \underline{(\underline{a}_n/\underline{b}_n)}_n,$$

в результате указанных действий над конечными десятичными дробями $\underline{a}_n$ и $\underline{b}_n$, имеющими не более чем n значащих цифр после запятой, получаются снова десятичные дроби с не более чем n значащими цифрами после запятой, при этом результат может быть получен с любой заданной степенью

точности. Именно таким образом и выполняют обычно действия с числами на практике.

З а м е ч а н и е 4. Отметим, что при построении способа записи действительных чисел последовательностями цифр за основу было взято число 10 (отрезки последовательно делились на десять равных частей). Вместо числа 10 можно взять любое натуральное число n. При использовании быстродействующих вычислительных машин часто употребляется так называемая двоичная система записи чисел, соответствующая случаю $n = 2$. При записи числа в двоичной системе участвуют только две цифры: 0 и 1. Например, число 14,625 в двоичной системе имеет вид 1110,101. Действительно, заметив, что $14 = 2^3 + 2^2$, $0{,}5 = 2^{-1}$, $0{,}125 = 2^{-3}$, получим

$$14{,}625 = 1 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 + 1 \cdot 2^{-1} + 0 \cdot 2^{-2} + 1 \cdot 2^{-3},$$

а цифры в двоичной системе записи числа являются соответствующими коэффициентами его разложения по степеням числа 2.

Аналогично теореме 6 доказывается, что между всеми действительными числами и всеми бесконечными двоичными дробями, не имеющими периода, состоящего только из одних единиц, существует взаимно однозначное соответствие.

Чтобы это доказать, надо последовательно делить отрезки $[n, n + 1]$, $n = 0, \pm 1, \pm 2, \dots$, не на десять частей, как это мы делали при доказательстве теоремы 6, а на два равных отрезка.

З а м е ч а н и е 5. При изложении теории действительных чисел можно рассуждать и в обратном порядке: определить действительные числа как бесконечные допустимые десятичные дроби и, используя эту запись, ввести в них соответствующим образом соотношение порядка и арифметические действия.

Существуют и другие способы построения действительных чисел, которые исходят из других конкретных объектов, однако все они приводят к совокупностям элементов, удовлетворяющих свойствам I—V п. 2.1. Напомним (см. п. 3.8*), что свойства I—V однозначно определяют совокупность элементов, обладающих этими свойствами. Однозначно в том смысле, что любые две совокупности, для элементов которых выполнены условия I—V, изоморфны относительно операций сложения, умножения и свойства упорядоченности. Здесь мы

встречаемся с характерной чертой математических методов исследования, для которых совершенно безразлична природа элементов, а важны лишь «количественные связи» между ними, которые в данном случае выражаются свойствами I—V.

4.11*. Счетные и несчетные множества

Возникает естественный вопрос: все ли бесконечные множества содержат одинаковое число элементов или бесконечности бывают разными? Прежде всего оказывается непонятным, что вообще означает термин «одинаковое число элементов» для бесконечных множеств. Сравнение бесконечных множеств по количеству содержащихся в них элементов, или, как принято говорить, по их мощности, удобно выполнять с помощью понятия взаимно однозначного соответствия между элементами множеств (см. п. 1.2*).

Определение 17. *Два множества называются равномощными, если между их элементами можно установить взаимно однозначное соответствие.*

С этой точки зрения натуральные числа $1, 2, \ldots, n, \ldots$ образуют множество, равномощное множеству четных чисел $2, 4, \ldots, 2n, \ldots$, хотя на первый взгляд последних кажется в два раза меньше. Требуемое взаимно однозначное соответствие получается, если натуральному числу n поставить в соответствие число $2n$, $n = 1, 2, \ldots$.

Четные числа составляют часть множества натуральных чисел, однако эти множества равномощны, следовательно, в случае бесконечных множеств часть может быть равна в нашем смысле целому!

Определение 18. *Множество, равномощное множеству всех натуральных чисел, называется счетным.*

Таким образом, если X счетно, то между множеством X и множеством натуральных чисел можно установить взаимно однозначное соответствие, или, как говорят, можно занумеровать элементы множества X, понимая под номером каждого элемента $x \in X$ соответствующее ему при указанном соответствии натуральное число.

Счетные множества являются в определенном смысле простейшими бесконечными множествами. Именно, справедлива следующая лемма.

ЛЕММА 1. *Любое бесконечное множество содержит счетное подмножество.*

Доказательство. Действительно, пусть X — бесконечное множество. Возьмем какой-либо его элемент и обозначим его x_1. В силу того, что X — бесконечное множество, в нем заведомо имеется хотя бы один элемент, отличный от элемента x_1. Выберем какой-либо из таких элементов и обозначим его x_2.

Пусть в множестве X уже выбраны элементы $x_1, \ldots, x_n$. Так как X — бесконечное множество, то в нем заведомо есть еще и другие элементы; выберем какой-либо из оставшихся элементов и обозначим его через x_{n+1} и т. д. В результате мы получили элементы $x_n \in X$, $n = 1, 2, \ldots$, которые образуют счетное подмножество множества X.

ЛЕММА 2. *Любое бесконечное подмножество счетного множества счетно.*

Доказательство. Пусть X — счетное множество, его элементы могут быть перенумерованы: $X = \{a_1, a_2, \ldots, a_n, \ldots\}$. Пусть Y — бесконечное подмножество множества X. Обозначим через b_1 первый встретившийся в ряде $a_1, a_2, \ldots, a_n, \ldots$ элемент множества Y, т. е. тот из элементов $a_n \in X$, который принадлежит множеству Y и имеет наименьший номер n_0: $b_1 = a_{n_0}$. Через b_2 обозначим тот из элементов a_n, который принадлежит множеству Y и имеет наименьший номер среди номеров $n > n_0$ и т. д. Каждый элемент множества Y имеется в ряде $a_1, a_2, \ldots, a_n, \ldots$, поэтому через какое-то конечное число шагов он будет обозначен через b_m, и так как множество Y бесконечно, то индекс m примет любое значение $1, 2, 3, \ldots$. Таким образом, все элементы множества Y окажутся занумерованными натуральными числами $m = 1, 2, \ldots$. Это и означает, что множество Y является счетным множеством. □

Следующая теорема дает интересный пример счетного множества.

ТЕОРЕМА 7. *Множество всех рациональных чисел счетно.*

Доказательство. Расположим рациональные числа в таблицу следующим способом. В первую строчку поместим все целые числа в порядке возрастания их абсолютной величины и так, что за каждым натуральным числом поставлено ему противоположное:

$$0, 1, -1, 2, \ldots, n, -n, \ldots, n \in \boldsymbol{N}.$$

Во вторую строчку поместим все несократимые рациональные дроби со знаменателем 2, упорядоченные по их абсолютной величине, причем снова за каждым положительным числом поставим ему противоположное:

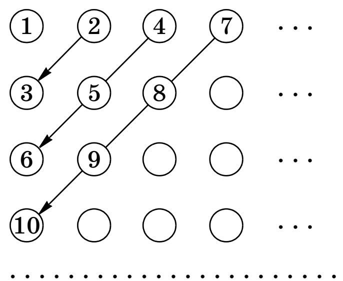
$$\frac{1}{2}, -\frac{1}{2}, \frac{3}{2}, -\frac{3}{2}, \frac{5}{2}, -\frac{5}{2} \dots .$$

Вообще, в n-ю строчку поместим все несократимые рациональные дроби со знаменателем n, упорядоченные по их абсолютной величине, и так, что за каждым положительным числом следует ему противоположное. В результате получим таблицу с бесконечным числом строк и столбцов:

$$\begin{array}{ccccc}
0 & -1 & 1 & 2 & -2\dots \\
\frac{1}{2} & -\frac{1}{2} & \frac{3}{2} & -\frac{3}{2} & \frac{5}{2}\dots \\
\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} & \frac{4}{3}\dots \\
\dots & \dots & \dots & \dots & \dots \\
\frac{1}{n} & -\frac{1}{n} & \dots & \dots & \dots \\
\dots & \dots & \dots & \dots & \dots
\end{array}$$

Очевидно, что каждое рациональное число попадает на какое-то место в этой таблице.

Занумеруем теперь элементы получившейся таблицы согласно следующей схеме (в кружочках стоят номера соответствующих элементов, стрелки указывают направление нумерации):



В результате все рациональные числа оказываются занумерованными, т. е. множество $\boldsymbol{Q}$ рациональных чисел счетно. □

Возникает естественный вопрос: а существуют ли бесконечные множества, не являющиеся счетными? Оказывается, что да, существуют, и они называются, естественно, *не-*

счетными множествами. Важный пример несчетных множеств устанавливается следующей ниже теоремой.

ТЕОРЕМА 8 (теорема Кантора). *Множество всех действительных чисел несчетно.*

Доказательство. Допустим противное: пусть удалось занумеровать все действительные числа: $x_1, x_2, \dots, x_n, \dots$; запишем их с помощью допустимых десятичных дробей:

$$\begin{aligned} x_1 &= \alpha_0^{(1)}, \alpha_1^{(1)}\alpha_2^{(1)} \dots \alpha_m^{(1)} \dots, \\ x_2 &= \alpha_0^{(2)}, \alpha_1^{(2)}\alpha_2^{(2)} \dots \alpha_m^{(2)} \dots, \\ &\dots\dots\dots\dots\dots\dots\dots \\ x_n &= \alpha_0^{(n)}, \alpha_1^{(n)}\alpha_2^{(n)} \dots \alpha_m^{(n)} \dots, \\ &\dots\dots\dots\dots\dots\dots\dots \end{aligned} \tag{4.60}$$

Здесь $\alpha_m^{(n)}$, $n = 1, 2, \dots, m = 1, 2, \dots$, обозначает одну из цифр $0, 1, 2, \dots, 9$, а $\alpha_0^{(n)}$, $n = 1, 2, \dots$, — целое число с тем или иным знаком. Выберем цифру α_n, $n = 1, 2, \dots$, так, чтобы $\alpha_n \neq \alpha_n^{(n)}$ и $\alpha_n \neq 9$. Тогда дробь $0, \alpha_1\alpha_2 \dots \alpha_n \dots$ является допустимой, но числа $a = 0, \alpha_1\alpha_2 \dots \alpha_n \dots$ заведомо нет среди чисел x_n, $n = 1, \dots, 2, \dots$, так как десятичная дробь $0, \alpha_1, \dots \dots \alpha_n \dots$ хотя бы одним десятичным знаком отличается от каждой из десятичных дробей (4.60). Полученное противоречие и доказывает теорему. □

СЛЕДСТВИЕ 1. *Множество действительных чисел, образующих любой интервал, несчетно.*

Доказательство. Покажем, что, более того, множество действительных чисел любого интервала равномощно множеству всех действительных чисел.

В самом деле, прежде всего любой интервал равномощен интервалу $(-1, +1)$. Взаимно однозначное соответствие между интервалом (a, b) и интервалом $(-1, +1)$ можно установить, например, с помощью линейного отображения $x = \frac{2t - a - b}{b - a}$. Если $a < t < b$, то $-1 < x < +1$. Интервал $(-1, +1)$ взаимно однозначно с помощью отображения

$$f(x) = \frac{x}{1 - |x|}, \quad -1 < x < 1,$$

отображается на всю действительную ось (проверьте это). Таким образом оказывается, что интервал (a, b) равномощен

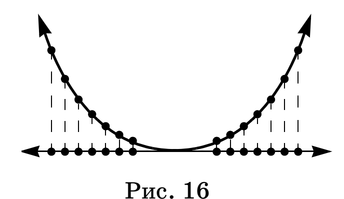


Рис. 16

Рис. 17

всей действительной оси и, следовательно, является несчетным множеством. □

Взаимно однозначное соответствие между интервалом и всей прямой можно наглядно осуществить геометрическим методом: сначала спроецируем открытую полуокружность, т. е. полуокружность без концевых точек, с помощью параллельной проекции на интервал (рис. 16) и тем самым установим между их точками взаимно однозначное соответствие. Затем с помощью центральной проекции из центра полуокружности спроецируем ее на прямую (рис. 17). Эта проекция также устанавливает взаимно однозначное соответствие, но на этот раз между полуокружностью и всей прямой.

СЛЕДСТВИЕ 2. *На любом интервале имеются иррациональные числа.*

Доказательство. Действительно, если бы на некотором интервале не оказалось иррационального числа, то это означало бы, что все точки этого интервала являются рациональными числами, т. е. подмножеством счетного множества рациональных чисел, и, значит, образуют конечное, или счетное, множество (см. лемму 2), что противоречит следствию 1. □

Замечание. В п. 4.10 доказано, что действительное число есть предел последовательности рациональных чисел (например, своих верхних десятичных приближений). Отсюда сразу следует, что всякий интервал содержит бесконечно много рациональных чисел. В самом деле, пусть задан интервал (a, b). Выберем какое-либо число $\xi \in (a, b)$, например $\xi = \frac{a+b}{2}$. Пусть для определенности $\xi > 0$. Тогда если $\bar{\xi}_n$, $n = 1, 2, \dots$, — верхние десятичные приближения для ξ, то из определения верхних десятичных приближений (см. п. 4.10) следует, что среди чисел $\bar{\xi}_n$ имеется бесконечно много раз-

личных между собой и что $\lim\limits_{n \to \infty} \bar{\xi}_n = \xi$. Интервал (a, b) является окрестностью выбранной точки ξ, поэтому, согласно определению предела последовательности, почти все рациональные числа $\bar{\xi}_n$ содержатся в этом интервале. Иначе говоря, найдется такой номер n_0, что для всех номеров $n > n_0$ будет выполняться неравенство $a < \bar{\xi}_n < b$, т. е. $\bar{\xi}_n$, $n = n_0 + 1, n_0 + 2, \ldots$, — искомые рациональные числа.

Таким образом, на любом интервале числовой оси содержатся как рациональные, так и иррациональные числа. Это свойство кратко выражают, говоря, что «рациональные и иррациональные числа образуют всюду плотные подмножества множества действительных чисел».

УПРАЖНЕНИЕ 17. Доказать, что множества точек интервала, отрезка и полуинтервала равномощны.

Метод, которым была доказана счетность рациональных чисел, позволяет доказать следующее предложение.

ТЕОРЕМА 9. *Сумма конечного или счетного множества непустых конечных или счетных множеств является конечным или счетным множеством.*

Доказательство. Обозначим через $\boldsymbol{N}_n$ множество номеров рациональных чисел, попавших в n-ю строку при нумерации рациональных чисел, изображенной на схеме на с. 143. Тогда $\boldsymbol{N} = \bigcup\limits_{n=1}^{\infty} \boldsymbol{N}_n$ является представлением множества натуральных чисел в виде объединения счетного множества попарно не пересекающихся счетных множеств.

Если X_n, $n = 1, 2, \ldots$, — конечные или счетные множества, то множества

$$Y_1 = X_1,\ Y_2 = X_2 \backslash X_1,\ \ldots,\ Y_n = X_n \backslash \bigcup_{k=1}^{n-1} X_k,\ \ldots$$

попарно не пересекаются и

$$X \stackrel{\text{def}}{=} \bigcup_{n=1}^{\infty} X_n = \bigcup_{n=1}^{\infty} Y_n.$$

Каждое из множеств X_n конечно или счетно, следовательно, и каждое множество Y_n конечно или счетно. Поэтому для каждого $n = 1, 2, \ldots$ существует взаимно однозначное

соответствие между элементами множества Y_n и некоторым подмножеством $A_n \subset \boldsymbol{N}_n$. В силу же того, что множества Y_n попарно не пересекаются, это соответствие порождает взаимно однозначное соответствие между множествами

$$X = \bigcup_{n=1}^{\infty} Y_n \quad \text{и} \quad \bigcup_{n=1}^{\infty} A_n \subset \boldsymbol{N}.$$

Согласно лемме 2, множество $\bigcup_{n=1}^{\infty} A_n$ конечно или счетно, как подмножество счетного множества. Следовательно, конечно или счетно множество X. □

Из этой теоремы вытекает следующее утверждение.

ТЕОРЕМА 10. *Множество всех конечных подмножеств счетного множества счетно.*

СЛЕДСТВИЕ. *Множество всевозможных конечных размещений из элементов счетного множества является счетным.*

Доказательство. Пусть X — счетное множество и $x_1, x_2, \ldots, x_n, \ldots$ — все его элементы. Очевидно, что множество всех его подмножеств, состоящих из одного элемента: $\{x_1\}, \{x_2\}, \ldots, \{x_n\}, \ldots$, является счетным множеством. Пусть уже доказано, что множество всех подмножеств множества X, состоящее из m элементов, $m \geqslant 1$, счетно. Обозначим эти подмножества $X_n^{(m)}$, $n = 1, 2, \ldots$. Если для каждого $n = 1, 2, \ldots$ к множеству $X_n^{(m)}$ поочередно добавлять по одному не принадлежащему ему элементу множества X, то получится счетное множество множеств, которые будем обозначать через $X_{n,k}^{(m)}$, $k = 1, 2, \ldots$. Совокупность всех множеств $X_{n,k}^{(m)}$, $n = 1, 2, \ldots, k = 1, 2, \ldots$, составляет совокупность всех подмножеств множества X, состоящих из $m + 1$ элементов. Этих множеств при фиксированном $n = 1, 2, \ldots$ счетное множество (индекс k принимает значения $1, 2, \ldots$), следовательно, всего счетное множество счетных множеств, объединение которых, согласно теореме 9, также является счетным множеством.

Итак, для любого $m = 1, 2, \ldots$ множество всех подмножеств множества X, состоящих из m элементов, счетно, а

тогда, согласно той же теореме 9, счетно и множество всех конечных подмножеств множества X. □

Докажем следствие. Чтобы получить всевозможные конечные размещения из элементов данного множества, надо во всех его конечных подмножествах сделать всевозможные перестановки элементов, которых для каждого такого подмножества конечное множество. Если исходное множество X было счетным, то, согласно теореме 10, множество всех его конечных подмножеств счетно, а поэтому множество всех конечных размещений из элементов множества X является счетной суммой конечных множеств и, следовательно, будет счетным. □

С помощью теоремы 10 легко устанавливается счетность множества всех алгебраических чисел, т. е. действительных чисел, являющихся корнями многочленов с целыми коэффициентами.

ТЕОРЕМА 11 (теорема Кантора). *Множество всех алгебраических чисел счетно.*

Доказательство. Каждый многочлен с целыми коэффициентами задается набором своих коэффициентов, т. е. в данном случае конечным множеством натуральных чисел, взятых в определенном порядке, т. е. некоторым конечным размещением натуральных чисел. Согласно следствию из теоремы 10, таких многочленов всего лишь счетное множество. Но каждый многочлен имеет лишь конечное множество действительных корней: их число не превышает степени многочлена. Следовательно, множество всех таких корней не более чем счетно (множества корней различных многочленов могут иметь непустое пересечение, поэтому нельзя еще утверждать счетность всего рассматриваемого множества корней). Однако каждое натуральное число является алгебраическим — оно является корнем уравнения $x - n = 0$. Поэтому алгебраические числа действительно составляют счетное множество. □

В теореме 10 не случайно рассматривались не все, а только конечные подмножества счетного множества. Для всех подмножеств справедливо другое утверждение.

ТЕОРЕМА 12. *Множество всех подмножеств счетного множества несчетно.*

Доказательство. Пусть $X = \{x_1, x_2, \dots, x_n, \dots\}$ — счетное множество. Поставим в соответствие каждому его подмножеству E функцию χ_E, определенную на множестве натуральных чисел и задаваемую следующим образом:

$$\chi_E(n) = \begin{cases} 1, \text{ если } x_n \in E, \\ 0, \text{ если } x_n \notin E \end{cases} \tag{4.61}$$

(функция χ_E называется *характеристической функцией множества* E). Это соответствие является взаимно однозначным между множеством всех подмножеств множества X и множеством всех функций, определенных на множестве натуральных чисел N и принимающих только значения 0 и 1. Каждой функции (4.61) поставим в соответствие действительное число, двоичная запись которого имеет вид

$$0, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n, \dots, \tag{4.62}$$

где $\varepsilon_n = 1$, если $\chi_E(n) = 1$, и $\varepsilon_n = 0$, если $\chi_E(n) = 0$. Это соответствие является взаимно однозначным соответствием между множеством всех функций (4.61) и множеством бесконечных двоичных дробей вида (4.62). Последнее множество содержит подмножество всевозможных дробей вида (4.62), не имеющих в периоде 1. Это подмножество находится во взаимно однозначном соответствии с несчетным множеством всех действительных чисел полуинтервала [0, 1). Поэтому множество двоичных дробей (4.62), как всякое множество, содержащее несчетное подмножество, также является несчетным. Следовательно, несчетным является и множество всех подмножеств счетного множества. □

Задача 6. Доказать, что для любого множества не существует взаимно однозначного соответствия между всеми его элементами и всеми его подмножествами.

4.12*. Верхний и нижний пределы последовательности

В п. 4.6 было показано, что любая числовая последовательность всегда имеет по крайней мере один частичный предел, конечный или бесконечный. Наибольший и наименьший из них (ниже будет показано, что они всегда существуют) игра-

ют особую роль в теории последовательностей. Здесь понятия «наибольший» и «наименьший» понимаются в смысле расширенного множества действительных чисел $\overline{\boldsymbol{R}}$ (см. п. 3.1), т. е., в частности, наибольшим (наименьшим) элементом множества $X \subset \overline{\boldsymbol{R}}$ может оказаться $+\infty$ (соответственно $-\infty$). Это имеет место тогда, когда $+\infty \in X$ ($-\infty \in X$). В данном случае это будет означать, что бесконечность соответствующего знака является частичным пределом рассматриваемой последовательности.

Не всякое множество на расширенной числовой прямой имеет наибольший (наименьший) элемент. Однако если это множество является множеством частичных пределов некоторой последовательности, то, как это будет показано ниже, в нем всегда существуют наибольший и наименьший элементы.

Определение 19. *Наибольший частичный предел последовательности* $\{x_n\}$ *называется ее верхним пределом и обозначается* $\overline{\lim\limits_{n\to\infty}}\, x_n$, *а наименьший частичный предел — нижним пределом и обозначается* $\underline{\lim\limits_{n\to\infty}}\, x_n$.

ТЕОРЕМА 13. *У любой последовательности* $\{x_n\}$ *существует как наибольший, так и наименьший частичный предел.*

Доказательство. Докажем существование наибольшего частичного предела. Для заданной последовательности $\{x_n\}$ возможны два случая: либо она ограничена сверху, либо нет. Если она не ограничена сверху, то положительная бесконечность $+\infty$ является ее частичным пределом и, очевидно, наибольшим, т. е. $\overline{\lim\limits_{n\to\infty}}\, x_n = +\infty$.

Если же последовательность $\{x_n\}$ ограничена сверху, то снова возможны два случая: либо множество ее конечных пределов, которое мы обозначим через A, не пусто, либо оно пусто. Рассмотрим сначала первый случай. Из ограниченности сверху данной последовательности $\{x_n\}$ следует и ограниченность сверху непустого множества A ее конечных частичных пределов. В силу этого, множество A имеет конечную верхнюю грань. Покажем, что $b = \sup A$ является частичным пределом, т. е. что $b \in A$. Действительно, если

бы $b \notin A$, то существовало бы такое $\varepsilon > 0$, что в интервале $(b - \varepsilon,\ b + \varepsilon)$ содержалось бы лишь конечное число членов последовательности $\{x_n\}$ (в частности, ни одного) и поэтому (почему?) в этом интервале не было бы ни одного элемента A, что противоречит условию $b = \sup A$.

Таким образом, $b \in A$ и, следовательно, является наибольшим элементом множества A, поэтому $b = \overline{\lim\limits_{n \to \infty}}\, x_n$.

В оставшемся случае, т. е. когда последовательность $\{x_n\}$ ограничена сверху и множество ее конечных частичных пределов A пусто, то $\lim\limits_{n \to \infty} x_n = -\infty$ (докажите это), т. е. в этом случае множество ее частичных пределов состоит из одного элемента $-\infty$, который тем самым является и наибольшим в этом множестве, т. е. здесь $\lim\limits_{n \to \infty} x_n = \lim\limits_{n \to \infty} x_n = -\infty$.

Аналогично для любой последовательности доказывается и существование наименьшего (конечного или бесконечного) частичного предела. □

УПРАЖНЕНИЕ 18. Пусть $x_n = \frac{(-1)^n}{n} + \frac{1 + (-1)^n}{2}$, $n = 1, 2, \ldots$. Найти $\underline{\lim\limits_{n \to \infty}}\, x_n$, $\overline{\lim\limits_{n \to \infty}}\, x_n$, $\inf\{x_n\}$, $\sup\{x_n\}$.

ТЕОРЕМА 14. *Для того чтобы число a было верхним пределом последовательности $\{x_n\}$, необходимо и достаточно выполнение для любого числа $\varepsilon > 0$ совокупности следующих двух условий.*

1. *Существует номер n_ε такой, что для всех номеров $n > n_\varepsilon$ справедливо неравенство $x_n < a + \varepsilon$.*

2. *Для любого номера n_0 существует номер n' (зависящий от ε и от n_0) такой, что $n' > n_0$ и $x_{n'} > a - \varepsilon$.*

Условие 1 означает, что при любом фиксированном $\varepsilon > 0$ в последовательности $\{x_n\}$ существует лишь конечное число членов x_n таких, что $x_n > a + \varepsilon$ (их номера не больше n_ε).

Условие же 2 означает, что при любом фиксированном $\varepsilon > 0$ в последовательности $\{x_n\}$ существует бесконечно много членов x_n таких, что $x_n > a - \varepsilon$.

Доказательство необходимости. Пусть $a = \varlimsup_{n \to \infty} x_n$ и $\varepsilon > 0$ фиксировано. Если бы на полуинтервале $[a + \varepsilon, +\infty)$ оказалось бесконечно много членов последовательности $\{x_n\}$, то нашлась бы подпоследовательность $\{x_n\}$, элементы которой принадлежат этому полуинтервалу и которая имеет конечный или бесконечный предел. Обозначим его через b. Очевидно, $b \geqslant a + \varepsilon > a$, что противоречит тому, что a — наибольший частичный предел последовательности. Условие 1 доказано.

Далее, верхний предел является и частичным пределом, следовательно, существует подпоследовательность $\{x_{n_k}\}$ такая, что $\lim\limits_{n \to \infty} x_{n_k} = a$. Почти все члены последовательности $\{x_{n_k}\}$ больше $a - \varepsilon$ и, следовательно, существует бесконечно много членов данной последовательности $\{x_n\}$, больших, чем $a - \varepsilon$. Условие 2 также доказано.

Доказательство достаточности. Пусть число a удовлетворяет условиям 1 и 2. Покажем, что тогда a является частичным пределом. Возьмем $\varepsilon = 1/k$, $k = 1, 2, \ldots$. Для каждого натурального k существует номер n_k такой, что $x_{n_k} > a - 1/k$ (согласно условию 2) и $x_{n_k} < a + 1/k$ (согласно условию 1). Так как для любого k множество элементов x_n данной последовательности, для которых выполняются неравенства $a - 1/k < x_n < a + 1/k$ бесконечно, то номера n_k можно последовательно ($k = 1, 2, \ldots$) выбрать так, чтобы $n_{k_1} < n_{k_2}$ при $k_1 < k_2$. В результате мы получим подпоследовательность $\{x_{n_k}\}$ данной последовательности $\{x_n\}$. Из неравенства $|a - x_{n_k}| < \frac{1}{k}$ следует, что $\lim\limits_{n \to \infty} x_{n_k} = a$, т. е. что a является частичным пределом последовательности $\{x_n\}$.

Покажем теперь, что число a является наибольшим частичным пределом. Действительно, если бы нашелся частичный предел b последовательности $\{x_n\}$ такой, что $b > a$, то,

выбрав $\varepsilon > 0$ так, что $a + \varepsilon < b$, получим, что на промежутке $(a + \varepsilon, +\infty)$ будет находиться бесконечно много членов последовательности $\{x_n\}$ (а именно почти все члены подпоследовательности, сходящейся к b). Это противоречит условию 1. □

УПРАЖНЕНИЯ. 19. Доказать, что, для того чтобы последовательность имела предел (конечный или бесконечный, равный одному из символов $+\infty$ или $-\infty$), необходимо и достаточно, чтобы $\overline{\lim\limits_{n\to\infty}} x_n = \underline{\lim\limits_{n\to\infty}} x_n$.

20. Доказать, что

$$\overline{\lim_{n\to\infty}}(x_n + y_n) \leqslant \overline{\lim_{n\to\infty}} x_n + \underline{\lim_{n\to\infty}} y_n.$$

21. Доказать, что

$$\overline{\lim_{n\to\infty}} x_n = \inf_n \{\sup_{m\geqslant n} x_m\} = \lim_{n\to\infty} \{\sup_{m\geqslant n} x_m\},$$

$$\underline{\lim_{n\to\infty}} x_n = \sup_n \{\inf_{m\geqslant n} x_m\} = \lim_{n\to\infty} \{\inf_{m\geqslant n} x_m\}.$$

§ 5.

Предел и непрерывность функций

5.1. Действительные функции

При изучении процессов реального мира (физических, химических, биологических, экономических и всевозможных других) мы постоянно встречаемся с характеризующими их величинами, меняющимися в течение рассматриваемых процессов. При этом часто бывает, что изменению одной величины сопутствует и изменение другой или даже, более того, изменение одной величины является причиной изменения другой. Взаимосвязные изменения числовых характеристик рассматриваемых величин приводят к их функциональной зависимости в соответствующих математических моделях. Поэтому понятие функции является одним из самых важных понятий в математике и ее приложениях.

В данном курсе математического анализа будут сначала изучаться только действительные функции одного дейст-

вительного аргумента, т. е. функции $f\colon X \to \boldsymbol{R}$, где $X \subset \boldsymbol{R}$ и $X \neq \varnothing$. Независимые и зависимые переменные называются в этом случае *действительными* (*вещественными*) переменными. Затем появятся функции многих переменных, т. е. функции, определенные на некотором множестве элементов, каждый из которых представляет собой упорядоченную совокупность чисел. Будут также изучаться функции, принимающие комплексные значения, функции, аргументами которых являются комплексные числа, и другие функции более общей природы.

Отметим, что функции, принимающие числовые значения (действительные или комплексные) называются *числовыми* или *скалярными*.

Над функциями, принимающими числовые значения, можно производить различные арифметические операции. Если даны две числовые функции f и g, определенные на одном и том же множестве X, а c — некоторое число (или, как часто говорят, постоянное), то функция cf определяется как функция, принимающая в каждой точке $x \in X$ значение $cf(x)$; функция $f + g$ — как функция, принимающая в каждой точке $x \in X$ значение $f(x) + g(x)$; fg — как функция, в каждой точке принимающая значение $f(x)g(x)$; наконец, f/g — как функция, в каждой точке $x \in X$ равная $f(x)/g(x)$ (что, конечно, имеет смысл лишь при $g(x) \neq 0$).

Числовая функция f, определенная на множестве X, называется ограниченной сверху (снизу), если множество ее значений ограничено сверху (снизу). Иначе говоря, функция f ограничена сверху (снизу), если существует такая постоянная M, что для каждого $x \in X$ выполняется неравенство $f(x) \leqslant M$ (соответственно $f(x) \geqslant M$).

Функция f, ограниченная на множестве X как сверху, так и снизу, называется просто *ограниченной* на этом множестве. Очевидно, что функция f ограничена на множестве X в том и только том случае, если существует такое число $M \geqslant 0$, что для каждого $x \in X$ выполняется неравенство $|f(x)| \leqslant M$.

Верхняя (нижняя) грань множества значений Y_f числовой функции $y = f(x)$, определенной на множестве X, называется *верхней* (*нижней*) *гранью* функции f и обозначается

$$\sup f, \quad \sup_X f, \quad \sup_{x \in X} f(x) \; \Big(\inf f, \quad \inf_X f, \quad \inf_{x \in X} f(x)\Big).$$

Более подробно это означает, что, например, $\lambda = \sup f$, если, во-первых, для каждого $x \in X$ выполняется неравенство $f(x) \leqslant \lambda$, и, во-вторых, для любого $\lambda' < \lambda$ существует такое $x_{\lambda'} \in X$, что $f(x_{\lambda'}) > \lambda'$. Индекс λ' у элемента множества X показывает, что он зависит от выбора числа λ'.

В приведенном определении верхняя (нижняя) грань функции может быть как конечной, так и бесконечной.

Согласно результатам п. 3.4, функция f ограничена сверху (снизу) на множестве X тогда и только тогда, когда она имеет на этом множестве конечную верхнюю (нижнюю) грань.

УПРАЖНЕНИЯ. *1.* Доказать, что если функция f не ограничена сверху (снизу) на отрезке $[a, b]$, то существует такая последовательность точек $x_n \in \in [a, b]$, $n = 1, 2, \ldots$, что $\lim\limits_{n \to \infty} f(x_n) = +\infty$ (соответственно $\lim\limits_{n \to \infty} f(x_n) = -\infty$).

2. Доказать, что если функция не ограничена на отрезке, то существует точка этого отрезка, в каждой окрестности которой функция не ограничена. Справедливо ли это утверждение для интервала?

3. Построить пример функции, определенной на отрезке и неограниченной на нем.

Будем говорить, что числовая функция f, определенная на множестве X, принимает в точке $x_0 \in X$ *наибольшее значение* (*наименьшее*), если $f(x) \leqslant f(x_0)$ (соответственно $f(x) \geqslant \geqslant f(x_0)$) для каждой точки $x \in X$. В этом случае будем писать $f(x_0) = \max\limits_X f$ или $f(x_0) = \max f$ (соответственно $f(x_0) = \min\limits_X f$ или $f(x_0) = \min f$).

Наибольшее (наименьшее) значение функции называется также ее *максимальным* (*минимальным*) значением. Максимальные и минимальные значения называются *экстремальными*.

Очевидно, что если функция f принимает в точке x_0 наибольшее (наименьшее) значение, то $f(x_0) = \sup f$ (соответственно $f(x_0) = \inf f$).

Отметим еще, что если заданы множества X, Y и соответствие f, ставящее в соответствие каждому элементу множества X единственный элемент множества Y, то этим функция f, определенная на множестве X и с множеством значений, содержащимся в множестве Y, полностью определена. В частности, безразлично, какой буквой обозначить аргумент и

какой — значение функции. Так, при заданном указанном соответствии f записи $y = f(x)$, $x \in X$, $y \in Y$ и $v = f(u)$, $u \in X$, $v \in Y$, обозначают одно и то же. Например, $y = \log_a x$, $x > 0$, и $x = \log_a y$, $y > 0$ обозначают одну и ту же функцию.

Сделаем в заключение еще одно замечание. В случае окрестности точки наряду с выражением «функция определена на окрестности» используется термин «функция определена в окрестности». В подобных выражениях предлоги «в» и «на» имеют одинаковый смысл.

5.2. Способы задания функций

В этой главе изучаются только действительные функции одной действительной переменной, т. е. такие функции f, что f: $X \to \boldsymbol{R}$, $X \subset \boldsymbol{R}$, $X \neq \varnothing$. Поэтому остановимся здесь только на способах задания таких функций.

Прежде всего функции могут задаваться с помощью формул: аналитический способ. Для этого используется некоторый запас изученных и специально обозначенных функций, алгебраические действия и предельный переход. Например, $y = ax + b$, $y = ax^2$, $y = \sin x$, $y = \sqrt{1 - x^2}$, $y = 1 + \sqrt{\lg \cos 2\pi x}$.

При этом всегда под функцией, заданной некоторой формулой, понимается функция, определенная на множестве всех тех действительных чисел, для которых, во-первых, указанная формула имеет смысл и, во-вторых, в процессе проведения всех необходимых вычислений по этой формуле получаются только действительные числа, причем окончательный результат вычислений для данного числа x из области определения рассматриваемой функции является ее *значением* в точке x. Так, областью существования функции $f(x) = \dfrac{x + |x|}{\sqrt{1 - x^2}}$ является интервал $(-1, 1)$, хотя эта функция принимает действительные значения и на полупрямой $x < -1$ (при всех $x \leqslant 0$ она равна нулю, но при $x < -1$ значение корня $\sqrt{1 - x^2}$ является существенно комплексным числом).

Отметим, что при таком определении действительные функции $f(x) = x$ и $f(x) = (\sqrt{x})^2$ имеют разные области определения: первая определена на множестве всех действитель-

ных чисел, а вторая — только на множестве всех неотрицательных.

Иногда функция задается с помощью нескольких формул, например

$$f(x) = \begin{cases} 2^x & \text{для } x > 0, \\ 0 & \text{для } x = 0, \\ x - 1 & \text{для } x < 0. \end{cases} \quad (5.1)$$

Функция может быть задана также просто с помощью описания соответствия. Поставим в соответствие каждому числу $x > 0$ число 1, числу 0 — число 0, а каждому $x < 0$ — число -1. В результате получим функцию, определенную на всей числовой оси и принимающую три значения: 1, 0 и -1. Эта функция имеет специальное обозначение[1] $\operatorname{sign} x$ и может быть записана с помощью нескольких формул:

$$\operatorname{sing} x = \begin{cases} 1 & \text{для } x > 0, \\ 0 & \text{для } x = 0, \\ -1 & \text{для } x < 0. \end{cases} \quad (5.2)$$

Другой пример: каждому рациональному числу поставим в соответствие число 1, а каждому иррациональному — число нуль. Полученная функция называется *функцией Дирихле*[2].

Отметим, что всякая формула является символической записью некоторого описанного ранее соответствия, так что, в конце концов, нет принципиального различия между заданием функции с помощью формулы или с помощью описания соответствия; это различие чисто внешнее.

Следует также иметь в виду, что всякая вновь определенная функция, если для нее ввести специальное обозначение, может служить для определения других функций с помощью формул, включающих этот новый символ.

Если речь идет о действительных функциях одного действительного аргумента, то для наглядного представления о характере функциональной зависимости часто строятся графики функций на координатной плоскости (координатной называется плоскость, на которой задана прямоугольная декартова система координат).

---

[1] Signum (лат.) означает «знак».

[2] Л. Дирихле (1805—1859) — немецкий математик.

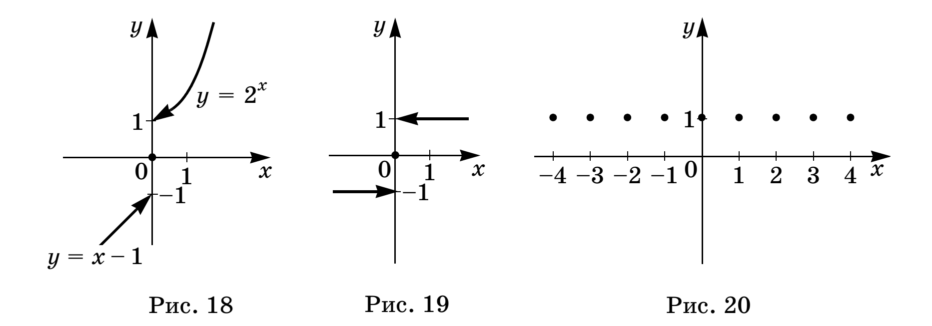


Рис. 18 Рис. 19 Рис. 20

Из общего определения графика функции (см. п. 1.2*) следует, что график функции $y = f(x)$ (x и y — числа, $x \in X$) представляет собой множество точек $(x,\ f(x))$, $x \in X$, на координатной плоскости переменных x и y.

Так, график функции (5.1) имеет вид, изображенный на рисунке 18, график функции $\operatorname{sign} x$ (см. формулы (5.2)) — на рисунке 19, а график функции $y = 1 + \sqrt{\lg \cos 2\pi x}$ состоит из отдельных точек, соответствующих целым значениям аргумента $x = 0, \pm 1, \pm 2, \ldots$, так как при остальных значениях аргумента выражение под знаком радикала принимает отрицательные значения (рис. 20).

Множество точек $\{(x,\ y)\colon x \in X,\ y \geqslant f(x)\}$ называется *надграфиком* данной функции f, а множество $\{(x,\ y)\colon x \in X,\ y \leqslant f(x)\}$ — ее *подграфиком*.

Графическое изображение функции также может служить для задания функциональной зависимости. Правда, это задание будет приближенно, потому что измерение отрезков практически можно производить лишь с определенной степенью точности. Примерами графического задания функций, встречающимися на практике, могут служить, например, показания осциллографа.

Функцию можно задать еще с помощью таблиц, т. е. для некоторых значений переменной x указать соответствующие значения переменной y. Данные таблиц могут быть получены как непосредственно из опыта, так и с помощью тех или иных математических расчетов. Примерами такого задания функций являются логарифмические таблицы тригонометрических функций. Само собой разумеется,

что функция, заданная с помощью таблицы, определена на конечном множестве точек.

Наконец, при проведении числовых расчетов на компьютерах функции задаются с помощью программ для их вычисления при нужных значениях аргумента или требуемые значения функции в готовом виде закладываются тем или иным способом в память компьютера.

Рассмотрим более подробно некоторые специальные аналитические способы задания функции.

Неявные функции. Пусть дано уравнение вида

$$F(x, y) = 0, \tag{5.3}$$

т. е. задана функция $F(x, y)$ двух действительных переменных x и y, и рассматриваются только такие пары x, y (если они существуют), для которых выполняется условие (5.3).

Пусть существует такое множество X, что для каждого $x_0 \in X$ существует по крайней мере одно число y, удовлетворяющее уравнению $F(x_0, y) = 0$. Обозначим одно из таких чисел через y_0 и поставим его в соответствие числу $x_0 \in X$. В результате получим функцию f, определенную на множестве X и такую, что $F(x_0, f(x_0)) = 0$ для всех $x_0 \in X$. В этом случае говорят, что функция f задается неявно уравнением (5.3). Одно и то же уравнение (5.3) задает, вообще говоря, не одну, а некоторое множество функций.

Функции, неявно задаваемые уравнениями вида (5.3), называются *неявными функциями* в отличие от функций, задаваемых формулой, разрешенной относительно переменной y, т. е. формулой вида $y = f(x)$.

Термин «неявная функция» отражает не характер функциональной зависимости, а лишь способ ее задания. Одна и та же функция может быть задана как явно, так и неявно. Например, функции $f_1(x) = \sqrt{1 - x^2}$ и $f_2(x) = -\sqrt{1 - x^2}$ могут быть заданы также и неявным образом с помощью уравнения $x^2 + y^2 - 1 = 0$ в том смысле, что они входят в совокупность функций, задаваемых этим уравнением.

Сложные функции. Напомним, что если заданы функции $y = f(x)$ и $z = F(y)$, причем область задания функции F содержит область значений функции f, то каждому x из области определения функции f естественным образом со-

ответствует z такое, что $z = F(y)$, где $y = f(x)$. Эта функция, определяемая соответствием $z = F[f(x)]$, называется, как известно, *сложной функцией* или *композицией* (*суперпозицией*) функции f и F и обозначается через $F \circ f$, т. е.

$$(F \circ f)(x) \overset{\text{def}}{=} F(f(x)).$$

Сложная функция отражает не характер функциональной зависимости, а лишь способ ее задания: может случиться, что одна и та же функция может быть задана как с помощью композиций каких-либо функций, так и без их помощи. Например, сложная функция $z = 2^y$, $y = \log_2 (1 + \sin^2 x)$, заданная с помощью суперпозиций показательной и логарифмической функций, может быть задана и без этой композиции: $z = 1 + \sin^2 x$.

Подобным образом можно рассматривать сложные функции, являющиеся композицией более чем двух функций, например функцию $w = \sin \lg \left(1 + \frac{1}{\sqrt{x}}\right)$ можно рассматривать как композицию следующих функций:

$$w = \sin v, \quad v = \lg u, \quad u = 1 + z, \quad z = \frac{1}{y}, \quad y = \sqrt{x}.$$

5.3. Элементарные функции и их классификация

Функции: постоянная $y = c$, c — константа, степенна́я $y = x^{\alpha}$, показательная $y = a^x$ $(a > 0)$, логарифмическая $y = \log_a x$ $(a > 0,\ a \neq 1)$, тригонометрические $y = \sin x$, $y = \cos x$, $y = \operatorname{tg} x$, $y = \operatorname{ctg} x$ и обратные тригонометрические $y = \arcsin x$, $y = \arccos x$, $y = \operatorname{arctg} x$, $y = \operatorname{arcctg} x$ — называются *основными элементарными функциями*.

Всякая функция, которая может быть явным образом задана с помощью формулы, содержащей лишь конечное число арифметических операций и композиций основных элементарных функций, называется элементарной функцией.

Под областью существования элементарной функции в соответствии с общим соглашением о функциях, заданных формулами (см. п. 5.2), обычно понимают множество всех действительных чисел x, для которых, во-первых, формула,

задающая рассматриваемую элементарную функцию, имеет смысл и, во-вторых, в процессе проведения всех необходимых вычислений по этой формуле получаются только действительные числа.

Выше рассмотренные функции, задаваемые формулами $y = ax + b$, $y = ax^2$, $y = \sqrt{1 - x^2}$, $y = 1 + \sqrt{\lg \cos 2\pi x}$, $y = \sin \ln \left(1 + \frac{1}{\sqrt{x}}\right)$, $y = \frac{x + |x|}{\sqrt{1 - x^2}}$ (заметим, что $|x| = \sqrt{x^2}$ — элементарная функция), являются элементарными функциями.

Элементарные функции обычно делят на следующие классы.

1. М н о г о ч л е н ы (п о л и н о м ы). К многочленам относятся функции, которые могут быть заданы формулами вида

$$y = P_n(x) = a_0 + a_1 x + \ldots + a_n x^n \equiv \sum_{k=0}^{n} a_k x^k.$$

Числа $a_0, a_1, \ldots, a_n$ называются *коэффициентами многочлена* $\boldsymbol{P}_n(x)$.

Если $a_n \neq 0$, то число n называется *степенью данного многочлена*. Многочлены первой степени называются также *линейными функциями*. Многочлен, у которого все коэффициенты равны нулю, называется *нулевым многочленом*.

Степень многочленов обладает тем свойством, что при перемножении ненулевых многочленов степень произведения равна сумме степеней сомножителей. Чтобы это свойство сохранялось и при умножении на нулевой многочлен, степень нулевого многочлена принимают равной минус бесконечности $-\infty$ и для любого действительного числа x по определению считают (см. п. 3.1), что его сумма с $-\infty$ равна $-\infty$: $-\infty + x = -\infty$, и что $(-\infty) + (-\infty) = -\infty$. В результате этих соглашений степень произведения многочленов, из которых по крайней мере один нулевой, оказывается равной $-\infty$, т. е. сумме степеней сомножителей.

2. Р а ц и о н а л ь н ы е ф у н к ц и и (р а ц и о н а л ь н ы е д р о б и). К этому классу относятся функции, которые могут быть заданы в виде

$$y = \frac{P(x)}{Q(x)},$$

где $P(x)$ и $Q(x)$ — многочлены, причем $Q(x)$ ненулевой многочлен.

Заметим, что класс многочленов содержится в классе рациональных функций.

3. Иррациональные функции. Иррациональной называется функция, не являющаяся рациональной, которая может быть задана с помощью композиций конечного числа рациональных функций, степенны́х функций с рациональными показателями и четырех арифметических действий. Например, функция $y = \sqrt[5]{(x-1)/(x^2+\sqrt{x})}$ является иррациональной функцией.

4. Трансцендентные функции. Элементарные функции, не являющиеся ни рациональными, ни иррациональными, называются трансцендентными элементарными функциями. Можно показать, что все прямые и обратные тригонометрические функции и показательная и логарифмическая функции являются трансцендентными функциями.

В данном курсе анализа изучаются в основном действительные функции от одного или нескольких действительных аргументов, поэтому вместо «действительная функция» будем говорить и писать просто «функция». В тех случаях, когда будут рассматриваться функции другой природы, это будет специально оговариваться или ясно из контекста.

5.4. Первое определение предела функции

Перейдем теперь к изучению одного из самых основных понятий математического анализа — понятию предела функции. Под «точками» будем понимать либо конечные точки, либо бесконечно удаленные, т. е. либо действительные числа, либо одну из бесконечностей ∞, $+\infty$ или $-\infty$. Дадим сначала определение предела функции $f: X \to \boldsymbol{R}$, $X \subset \boldsymbol{R}$ в терминах пределов последовательностей. Это определение часто называют определением предела функции по Гейне[1].

Определение 1. *Точка a называется пределом функции $f: X \to \boldsymbol{R}$ в точке x_0 (или, что то же, при $x \to x_0$[2]), если*

---

[1] Г. Гейне (1821—1881) — немецкий математик.

[2] Запись «при $x \to x_0$» читается: «при x, стремящемся к x_0».

для любой последовательности $x_n \in X$, $n = 1, 2, \ldots$, *имеющей своим пределом точку* x_0, *т. е. такой, что*

$$\lim_{n \to \infty} x_n = x_0, \qquad (5.4)$$

последовательность $\{f(x_n)\}$ *имеет своим пределом точку* a, *т. е.*

$$\lim_{n \to \infty} f(x_n) = a. \qquad (5.5)$$

В том случае, когда a является пределом функции f в точке x_0, пишут

$$\lim_{x \to x_0} f(x) = a \text{ или } f(x) \to a \text{ при } x \to x_0,$$

а если x_0 — число: $x_0 \in \boldsymbol{R}$, то иногда пишут также

$$\lim_{x - x_0 \to 0} f(x) = a.$$

В символической форме определение предела функции записывается следующим образом:

$$\lim_{x \to x_0} f(x) = a \overset{\text{def}}{\Leftrightarrow}$$

$$\overset{\text{def}}{\Leftrightarrow} \{\forall\, x_n \in X, n = 1, 2, \ldots ,\ \lim_{x \to \infty} x_n = x_0 \Rightarrow \lim_{n \to \infty} f(x_n) = a\}.$$

Подчеркнем, что в определении 1 x_0 и a могут быть как действительными числами, так и бесконечностями: ∞, $+\infty$ и $-\infty$. Если $\lim\limits_{x \to x_0} f(x) = a$ и a — действительное число, то говорят, что в точке x_0 функция f имеет конечный предел (равный a).

Определение предела при заданной функции $f : X \to \boldsymbol{R}$ содержательно, конечно, только тогда, когда для точки x_0 действительно существуют последовательности точек $x_n \in X$, $n = 1, 2, \ldots$, имеющие своим пределом (конечным или бесконечным) точку x_0: $\lim\limits_{n \to \infty} x_n = x_0$.

Определение 2. *Пусть* $X \subset \boldsymbol{R}$. *Точка* x_0, *для которой существует последовательность* $x_n \in X$, $n = 1, 2, \ldots$, *имеющая своим пределом точку* x_0,

$$\lim_{n \to \infty} x_n = x_0 \qquad (5.6)$$

называется точкой прикосновения множества X.

Если точка прикосновения x_0 множества X является одной из бесконечностей ∞, $+\infty$ или $-\infty$, то она называется также и *бесконечно удаленной точкой прикосновения*. Очевидно, что если $x_0 = \infty$ — бесконечно удаленная точка прикосновения множества X, то оно не ограничено, если $x_0 = +\infty$ ($x_0 = -\infty$) является бесконечно удаленной точкой прикосновения множества X, то оно не ограничено сверху (снизу).

Очевидно, что любая точка x_0, принадлежащая самому множеству X, является его точкой прикосновения, так как стационарная последовательность $x_n = x_0 \in X$, $n = 1, 2, \ldots$, удовлетворяет условиям определения 2: $\lim\limits_{n \to \infty} x_n = x_0$ и $x_n \in X$, $n = 1, 2, \ldots$. Но, безусловно, у множеств могут существовать и конечные точки прикосновения, не принадлежащие этим множествам. Так, например, точки $x = a$ и $x = b$ являются точками прикосновения интервала (a, b) и не содержатся в нем.

УПРАЖНЕНИЕ 4. Доказать, что если точка x_0 является точкой прикосновения множества X и $X \subset Y \subset \boldsymbol{R}$, то точка x_0 является и точкой прикосновения множества Y.

З а м е ч а н и е 1. Нетрудно убедиться, что некоторая точка является точкой прикосновения данного множества тогда и только тогда, когда любая ее окрестность пересекается с этим множеством.

Действительно, если x_0 — точка прикосновения множества X, то существует такая последовательность $x_n \in X$, $n = 1, 2, \ldots$, что $\lim\limits_{n \to \infty} x_n = x_0$ и, следовательно, в любую окрестность точки x_0 попадут все члены этой последовательности начиная с некоторого номера (а они являются точками множества X).

Наоборот, если в любой окрестности точки x_0 имеются точки множества X, то, выбрав для каждого натурального n какую-либо точку в непустом по условию пересечения $X \cap U\left(x_0, \frac{1}{n}\right)$ и обозначив ее через x_n, т. е.

$$x_n \in X \cap U\left(x_0, \frac{1}{n}\right), \; n = 1, 2, \ldots,$$

получим такую последовательность $\{x_n\}$, что $x_n \to x_0$ (см. замечание 1 в п. 4.2) и $x_n \in X$, $n = 1, 2, \ldots$. Это и означает, что x_0 является точкой прикосновения множества X. □

Для любого непустого множества $X \subset \boldsymbol{R}$ его верхняя грань $\beta = \sup X$ и нижняя грань $\alpha = \inf X$ являются его точками прикосновения (они могут быть конечными или бесконечными). Это сразу следует, в силу леммы 1, из определения 4′ верхней грани и определения 5′ нижней грани множества, так как в этих определениях требуется, чтобы в любой окрестности соответствующей грани множества находилась его точка (и даже с одной стороны от рассматриваемой грани).

Из определения 1 предела функции непосредственно следует, что в точке прикосновения своего множества определения функция не может иметь двух различных пределов, т. е. указанное определение однозначно.

Далее из определения предела функции следует, что значения, которые принимает функция в точках, лежащих вне любой фиксированной окрестности точки x_0, не влияют ни на существование, ни на значение предела функции в точке x_0. Образно говоря, существует или нет предел функции в данной точке x_0, а если существует, то каково его значение, полностью определяется значениями функции на пересечении $U(x_0) \cap X$ любой окрестности $U(x_0)$ точки x_0 (являющейся точкой прикосновения множества X) с самим этим множеством.

Действительно, какова бы ни была окрестность $U(x_0)$ и какова бы ни была последовательность $x_n \in X$, $n = 1, 2, \ldots$, $\lim\limits_{n \to \infty} x_n = x_0$, найдется такой номер n_0, что при $n > n_0$, $n \in \boldsymbol{N}$, будет иметь место включение $x_n \in U(x_0) \cap X$, а конечное число членов $f(x_1)$, $f(x_2)$, ... , $f(x_{n_0})$ последовательности $\{f(x_n)\}$ не влияет ни на существование ее предела, ни на его значение, если он существует.

Свойства функции, которые зависят лишь от значений функции в любых окрестностях рассматриваемой точки, точнее говоря, которые не меняются при переходе к сужению функции на пересечение ее множества определения с любой окрестностью точки, называются *локальными свой-*

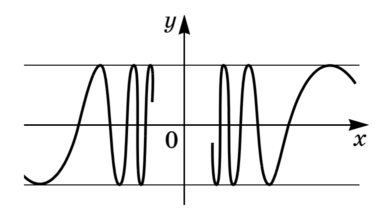


Рис. 21

ствами функции в данной точке. Из сказанного следует, что как существование предела в точке, так и его значение, если он существует, являются локальными свойствами функции в этой точке.

Примеры. 1. Пусть

$$f(x) = \frac{2x^2 + x - 1}{x - 1}. \tag{5.7}$$

Множество X, на котором определена функция (5.7), получается из множества всех действительных чисел $\boldsymbol{R}$ удалением из него единицы: $X = \boldsymbol{R}\backslash\{1\}$. Выясним, существует или нет предел функции f (5.7) в точке $x_0 = 0$.

Возьмем какую-либо последовательность $x_n \in X$, $n = 1, 2, \ldots$, такую, что $\lim\limits_{n \to \infty} x_n = 0$. Тогда на основании теорем п. 4.9 получаем

$$\lim_{n \to \infty} f(x_n) = \lim_{n \to \infty} \frac{2x_n^2 + x_n - 1}{x_n - 1} = \frac{2(\lim\limits_{n \to \infty} x_n)^2 + \lim\limits_{n \to \infty} x_n - 1}{\lim\limits_{n \to \infty} x_n - 1} = 1.$$

Таким образом, существует $\lim\limits_{n \to \infty} f(x_n) = 1$, а так как он не зависит от выбора последовательности $x_n \to 0$, $x_n \in X$, $n = 1, 2, \ldots$, то существует и предел $\lim\limits_{n \to \infty} f(x) = 1$.

2. Рассмотрим (рис. 21) функцию

$$f(x) = \sin\frac{1}{x}. \tag{5.8}$$

Она определена на множестве $X = \boldsymbol{R}\backslash\{0\}$. Снова выясним, существует или нет у функции f предел в точке $x_0 = 0$. Возьмем две последовательности

$$x_n = \frac{1}{\pi n} \text{ и } x'_n = \frac{1}{\pi/2 + 2\pi n}, \quad n = 1, 2, \ldots .$$

Очевидно, что $\lim\limits_{n \to \infty} x_n = \lim\limits_{n \to \infty} x'_n = 0$, $x_n \neq 0$, $x'_n \neq 0$ (условие $x \neq 0$ в данном случае означает, что $x \in X$), $f(x_n) = \sin \pi n = 0$, $f(x'_n) = \sin\left(\frac{\pi}{2} + 2\pi n\right) = 1$, $n = 1, 2, \ldots$. Поэтому

$\lim\limits_{n \to \infty} f(x_n) = 0$ и $\lim\limits_{n \to \infty} f(x'_n) = 1$, а это означает, что предела функции (5.8) при $x \to 0$ не существует.

3. Пусть

$$f(x) = \frac{x^2 + x + 1}{x^2 - 2}.$$

Найдем предел этой функции при $x \to \infty$. Ее областью определения является множество $X = \boldsymbol{R} \setminus \{\sqrt{2}, -\sqrt{2}\}$. Взяв какую-либо последовательность $x_n \in X$, $n = 1, 2, \ldots$, $\lim\limits_{n \to \infty} x_n = \infty$, будем иметь

$$\lim_{n \to \infty} f(x_n) = \lim_{n \to \infty} \frac{x_n^2 + x_n + 1}{x_n^2 - 2} =$$

$$= \lim_{n \to \infty} \frac{1 + \frac{1}{x_n} + \frac{1}{x_n^2}}{1 - \frac{2}{x_n^2}} = \frac{1 + \lim\limits_{n\to\infty} \frac{1}{x_n} + \lim\limits_{n\to\infty} \frac{1}{x_n^2}}{1 - 2\lim\limits_{n\to\infty} \frac{1}{x_n^2}} = 1.$$

Отсюда следует, что $\lim\limits_{n \to \infty} \dfrac{x^2 + x + 1}{x^2 - 2} = 1$.

УПРАЖНЕНИЕ 5. Доказать, что предел $\lim\limits_{n \to \infty} \dfrac{x}{\sqrt{x^2 + 1}}$ не существует, а пределы $\lim\limits_{n \to +\infty} \dfrac{x}{\sqrt{x^2 + 1}}$ и $\lim\limits_{n \to -\infty} \dfrac{x}{\sqrt{x^2 + 1}}$ существуют, и найти их.

При рассмотрении пределов функции часто приходится иметь дело с пределами сужений функций на том или ином множестве, т. е. с пределами функций: получающихся из данных функций, рассмотрением их не на всем множестве, на котором они заданы, а на каком-то содержащемся в нем.

Определение 3. *Пусть $f : X \to \boldsymbol{R}$. Предел в точке x_0 сужения $f_E : E \to \boldsymbol{R}$, $E \subset X$, функции f на множество E называется пределом функции f по множеству E в этой точке и обозначается через*

$$\lim_{\substack{x \to x_0 \\ x \in E}} f(x).$$

Таким образом,

$$\lim_{\substack{x \to x_0 \\ x \in E}} f(x) \overset{\text{def}}{=} \lim_{x \to x_0} f_E(x), \tag{5.9}$$

т. е. предел функции по множеству $E \subset X$ не является новым понятием по сравнению с пределом функции — это просто предел в смысле определения 1 функции, являющейся сужением данной на некоторое множество E.

Иногда при обозначении предела по множеству E вместо указания, что $x \in E$, употребляются для краткости другие обозначения, смысл которых бывает ясен из контекста. Например, при $E = X \setminus \{x_0\}$ пишут $\lim\limits_{x \to x_0,\, x \neq x_0} f(x)$.

Понятие предела функции по множеству в точке x_0 содержательно только для такого множества E, для которого точка x_0 является его точкой прикосновения (в этом случае она является, очевидно, и точкой прикосновения множества X).

Употребляя терминологию определения 3, можно сказать, что предел функции в смысле определения 1 является ее пределом в точке x_0 по всему множеству определения X функции f:

$$\lim_{x \to x_0} f(x) = \lim_{\substack{x \to x_0 \\ x \in X}} f(x) .$$

В том случае, если функция f задана некоторой формулой, под $\lim\limits_{x \to x_0} f(x)$ понимается предел этой функции в точке x_0 по всему множеству значений X, для которых указанная формула имеет смысл и для которых в процессе проведения всех вычислений по этой формуле получаются только действительные числа (см. п. 5.2).

Пример 4. Пусть f — функция Дирихле (см. п. 5.2), т. е. функция, равная 1 на множестве $\boldsymbol{Q}$ всех рациональных чисел и нулю на множестве I всех иррациональных чисел. Тогда в точке $x_0 = 0$ ее предел по множеству рациональных чисел равен 1:

$$\lim_{\substack{x \to 0 \\ x \in \boldsymbol{Q}}} f(x) = 1,$$

а по множеству иррациональных чисел — нулю:

$$\lim_{\substack{x \to 0 \\ x \in I}} f(x) = 0.$$

По всему же множеству действительных чисел (т. е. по множеству определения функции Дирихле) предел ее в точке $x_0 = 0$ не существует, так как уже существование или нет предела последовательности $\{f(x_n)\}$ при $n \to \infty$ зависит в данном случае от выбора последовательности $\{x_n\}$, стремящейся к нулю.

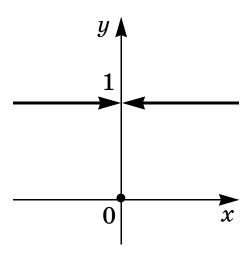


Рис. 22

Отметим следующее простое утверждение.

ЛЕММА 1. *Если $f : X \to \boldsymbol{R}$, $E \subset X$, x_0 — точка прикосновения множества E и существует предел $\lim\limits_{x \to x_0} f(x)$ функции f в точке x_0 (т. е. предел по множеству X), то в этой точке существует и предел функции f по множеству E и значения обоих пределов равны*:

$$\lim_{\substack{x \to x_0 \\ x \in E}} f(x) = \lim_{x \to x_0} f(x). \tag{5.10}$$

Доказательство. Если для любой последовательности $x_n \in X$, $n = 1, 2, \ldots$, $\lim\limits_{n \to \infty} x_n = x_0$ все последовательности $\{f(x_n)\}$ имеют один и тот же предел a, то это заведомо верно и для любой последовательности $x_n \in E$, $n = 1, 2, \ldots$, $\lim\limits_{n \to x_0} x_n = x_0$, так как $E \subset X$. □

Отметим один часто встречающийся случай предела функции в точке, когда предел берется по проколотой окрестности (она определена ниже) этой точки или по пересечению проколотой окрестности с множеством определения рассматриваемой функции.

Определение 4. *Проколотой окрестностью точки x_0 называется множество, получающееся удалением точки x_0 из ее окрестности.*

Проколотая окрестность точки x_0 обозначается через $\mathring{U}(x_0)$:

$$\mathring{U}(x_0) \overset{\text{def}}{=} U(x_0) \setminus \{x_0\}. \tag{5.11}$$

В частности, проколотая ε-окрестность точки x_0 обозначает $\mathring{U}(x_0, \varepsilon)$.

Пример 5. Рассмотрим функцию $f(x) = |\text{sign } x|$ (определение функции $\text{sign } x$ см в п. 5.2). Ее график изображен на рисунке 22. Какова бы ни была окрестность нуля $U(0)$, у этой функции в точке $x_0 = 0$, очевидно, существует предел по проколотой окрестности $\overset{\circ}{U}(0)$:

$$\lim_{\substack{x \to 0 \\ x \in \overset{\circ}{U}(0)}} |\text{sign} x| = 1.$$

Вместе с тем предел $\lim\limits_{\substack{x \to 0 \\ x \in U(0)}} |\text{sign} x|$ по всей окрестности $U(0)$ в точке $x_0 = 0$ у функции $|\text{sing } x|$ не существует, так как, например, для последовательности

$$x_n = \begin{cases} \frac{1}{n}, \text{ если } n = 2k, \\ 0, \text{ если } n = 2k - 1, \end{cases} \quad k = 1, 2, \dots ,$$

имеем $\lim\limits_{n \to \infty} x_n = 0$ (и, следовательно, все ее члены начиная с некоторого будут лежать в заданной окрестности $U(0)$), а последовательность $|\text{sign } x_n|$ не имеет предела (на четных местах у нее стоят единицы, а на нечетных — нули).

Рассмотренные примеры 4 и 5 показывают, в частности, что одна и та же функция может по одному множеству иметь предел в некоторой точке, а по другому не иметь предела в той же точке или иметь, но другой.

Замечание 2. Если функции f и g определены в некоторой окрестности $U(x_0)$ точки x_0, кроме, быть может, самой точки x_0, и при каждом x, принадлежащем проколотой окрестности $\overset{\circ}{U}(x_0)$: $x \in U(x_0)$, они принимают равные значения $f(x) = g(x)$, то тогда их пределы по проколотой окрестности $\overset{\circ}{U}(x_0)$ одновременно существуют или нет, а если существуют, то равны между собой:

$$\lim_{\substack{x \to x_0 \\ x \in \overset{\circ}{U}(x_0)}} f(x) = \lim_{\substack{x \to x_0 \\ x \in \overset{\circ}{U}(x_0)}} g(x),$$

ибо в их определении участвуют только значения функций в точках проколотой окрестности $\overset{\circ}{U}(x_0)$.

На этом простом замечании основано *правило раскрытия неопределенностей вида* $\frac{0}{0}$ *с помощью сокращения дробей*. Поясним его на примере.

Пример 6. Найдем

$$\lim_{x \to 0} \frac{(2x^2 + x - 1)x}{x^2 - x}. \tag{5.12}$$

Повторяя рассуждения, аналогичные тем, с помощью которых был вычислен предел в примере 1, приходим к выражению $\frac{0}{0}$, т. е. к неопределенности, и тем самым не получаем ответа ни на вопрос о существовании предела (5.12), ни на вопрос о его значении, если он существует. Поэтому рассмотрим функцию

$$f(x) = \frac{2x^2 + x - 1}{x - 1},$$

получающуюся из функции

$$g(x) = \frac{(2x^2 + x - 1)x}{x^2 - x},$$

стоящей под знаком предела в выражении (5.12), сокращением правой части равенства на x. Функции f и g совпадают в проколотой окрестности $\mathring{U}(0, 1) = (-1,1)\setminus\{0\}$ точки $x_0 = 0$ и поэтому, согласно сделанному выше замечанию, одновременно имеют или нет пределы в этой точке по указанной проколотой окрестности, причем в случае существования этих пределов они равны. В примере же 1 было показано, что $\lim\limits_{x \to 0} f(x) = 1$ по всей области определения функции f, следовательно, и по ее подмножеству $\mathring{U}(0, 1)$. Таким образом,

$$\lim_{x \to 0} g(x) = \lim_{\substack{x \to x_0 \\ x \in \mathring{U}(0,1)}} g(x) = \lim_{\substack{x \to x_0 \\ x \in \mathring{U}(0,1)}} f(x) = \lim_{x \to 0} f(x) = 1$$

(первое равенство справедливо в силу того, что предел является локальным свойством функции). Эти рассуждения являются обоснованием вычислений, которые в обычно употребляемой записи имеют следующий вид:

$$\lim_{x \to 0} \frac{(2x^2 + x - 1)x}{x^2 - 1} = \lim_{x \to 0} \frac{2x^2 + x - 1}{x - 1} = 1.$$

З а м е ч а н и е 3. Частным случаем предела функции (конечного или бесконечного) является предел последовательности (конечный или бесконечный). Действительно, последовательность $x_n \in \boldsymbol{R}$, $n = 1, 2, \dots$, является функцией, определенной на множестве натуральных чисел: $f : \boldsymbol{N} \to \boldsymbol{R}$, причем $f(n) = x_n$, $n = 1, 2, \dots$. Данное раньше определение предела последовательности $\lim\limits_{n \to \infty} x_n$ (см. определение 1 в п. 4.1) и определение ее предела как частного случая определения предела функции $\lim\limits_{n \to \infty} f(n)$ (см. определение 1 этого пункта) равносильны в силу того, что если последовательность $\{x_n\}$ имеет предел (конечный или бесконечный) в смысле определения 1 п. 4.1, то при любом выборе последовательности натуральных чисел $\{n_k\}$ такой, что $\lim\limits_{n \to \infty} n_k = +\infty$, последовательность $\{x_{n_k}\}$ имеет тот же предел, что и последовательность $\{x_n\}$ (см. лемму в п. 4.3), как это и требуется в определении 1.

5.5. Непрерывные функции

При рассмотрении предела функции $f : X \to \boldsymbol{R}$ при $x \to x_0$ может оказаться, что $x_0 \in X$ (тогда x_0 является числом: $x_0 \in \boldsymbol{R}$) или, наоборот, что $x_0 \notin X$. Случай $x_0 \in X$ представляет особый интерес, так как приводит к важному понятию непрерывной функции. Его изучение начнем с доказательства следующей леммы.

ЛЕММА 2. *Пусть $f : X \to \boldsymbol{R}$ и $x_0 \in X$. Тогда, для того чтобы функция f имела предел в точке x_0, необходимо и достаточно, чтобы*

$$\lim_{x \to x_0} f(x) = f(x_0). \tag{5.13}$$

Д о к а з а т е л ь с т в о. Достаточность условия (5.13) для существования предела функции f в точке x_0 очевидна: это условие даже сильнее, так как оно утверждает не только существование предела, но и определяет его значение, равное $f(x_0)$.

Докажем необходимость условия (5.13) для существования предела функции f в точке x_0. Пусть у функции f в точке x_0 существует предел, равный a:

$$\lim_{x \to x_0} f(x) = a. \tag{5.14}$$

Согласно определению предела, это означает, что для любой последовательности $x_n \in X$, $n = 1, 2, \ldots$, $\lim\limits_{n \to \infty} x_n = x_0$, справедливо равенство

$$\lim_{n \to \infty} f(x_n) = a. \tag{5.15}$$

В частности, поскольку $x_0 \in X$, это равенство справедливо и для стационарной последовательности, составленной из одной точки x_0, т. е. для последовательности $x_n = x_0$, $n = 1, 2, \ldots$. В этом случае (5.15) имеет вид

$$\lim_{n \to \infty} f(x_0) = a. \tag{5.16}$$

С другой стороны, так как предел постоянной равен самой этой постоянной, то

$$\lim_{n \to \infty} f(x_0) = f(x_0). \tag{5.17}$$

Сравнив (5.16) и (5.17), получим $f(x_0) = a$. □

Определим теперь понятие функции, непрерывной в данной точке.

Определение 5. *Функция $f : X \to \boldsymbol{R}$ называется непрерывной в точке $x_0 \in X$, если*

$$\lim_{x \to x_0} f(x) = f(x_0). \tag{5.18}$$

Условие (5.18) означает, что в случае непрерывности функции в точке x_0 предел f в этой точке находится по очень простому правилу: следует вычислить значение самой функции f в точке x_0.

Согласно лемме 2, условие (5.18) равносильно тому, что функция $f : X \to \boldsymbol{R}$ имеет предел в точке x_0 и что $x_0 \in X$. Само собой разумеется, что в том случае, когда для функции $f : X \to \boldsymbol{R}$ предел $\lim\limits_{x \to x_0} f(x)$ равен одной из бесконечностей ∞, $+\infty$ или $-\infty$, заведомо $x_0 \notin X$. В противном случае для

стационарной последовательности $x_n = x_0$, $n = 1, 2, \dots$, имело бы место $\lim\limits_{n \to \infty} f(x_n) = \lim\limits_{n \to \infty} f(x_0) = f(x_0)$, и так как, по условию, функция f принимает только числовые значения, то вопреки предположению предел $\lim\limits_{x \to x_0} f(x)$ был бы конечным. Из сказанного следует, в частности, что если у функции в некоторой точке существует бесконечный предел, то в ней функция заведомо не является непрерывной.

Отметим еще, что в определении 5 точка x_0, в которой определяется понятие непрерывности функции, принадлежит числовой прямой $\boldsymbol{R}$, т. е. не является бесконечно удаленной точкой.

Для проведения анализа понятия непрерывности функции в точке дадим определения изолированных и предельных точек множеств.

Определение 6. *Точка $x_0 \in X$ называется изолированной точкой множества $X \subset \boldsymbol{R}$, если существует окрестность $U(x_0)$ этой точки, пересечение $U(x_0) \cap X$ которой с множеством X состоит только из одной точки x_0:*

$$U(x_0) \cap X = \{x_0\}. \qquad (5.19)$$

Все точки множества натуральных чисел $\boldsymbol{N}$ изолированы, а множество $\boldsymbol{Q}$ всех рациональных чисел вовсе не имеет изолированных точек.

Определение 7. *Точка $x_0 \in \boldsymbol{R}$ называется предельной точкой множества $X \subset \boldsymbol{R}$, если в любой ее окрестности существует отличная от нее точка, принадлежащая множеству X.*

Иначе говоря, *точка x_0 называется предельной точкой множества X, если всякая ее проколотая окрестность имеет с этим множеством непустое пересечение: для любой проколотой окрестности $\mathring{U}(x_0)$ выполняется условие $\mathring{U}(x_0) \cap X \neq \varnothing$.*

Предельная точка множества может как принадлежать самому множеству, так и не принадлежать. Например, каждая точка отрезка $[a, b]$ является предельной точкой интервала (a, b). При этом точки a и b не принадлежат указанному интервалу, а все остальные содержатся в нем.

Если точка принадлежит множеству, то, в силу определений 6 и 7, она является его изолированной точкой тогда и только тогда, когда не является его предельной точкой.

Всякая точка прикосновения x_0 множества является либо изолированной точкой этого множества, либо его предельной точкой, так как либо в любой окрестности содержится точка множества, отличная от x_0 (тогда x_0 — предельная), либо у x_0 существует окрестность, не содержащая не совпадающих с x_0 точек множества, и тогда, в результате того что в этой окрестности все-таки существует точка множества (поскольку точка x_0 по условию является его точкой прикосновения), этой точкой оказывается x_0; следовательно, во-первых, x_0 принадлежит рассматриваемому множеству, а во-вторых, является его изолированной точкой.

Справедливо следующее предложение.

ЛЕММА 3. *Всякая функция непрерывна в каждой изолированной точке множества своего определения.*

Доказательство. Пусть x_0 — изолированная точка множества определения X функции f. Тогда, согласно определению 6, существует окрестность $U(x_0)$ точки x_0, пересечение которой с множеством X состоит из единственной точки x_0, т. е. $U(x_0) \cap X = \{x_0\}$. Какова бы ни была последовательность $x_n \in X$, $n = 1, 2, \dots$, $\lim\limits_{n \to \infty} x_n = x_0$, для указанной окрестности, в силу определения предела последовательности, существует такой номер n_0, что для всех номеров $n > n_0$ выполняется включение $x_n \in U(x_0)$ и, следовательно, включение $x_n \in U(x_0) \cap X$. Но $U(x_0) \cap X = \{x_0\}$, поэтому для всех $n > n_0$ имеем $x_n = x_0$. Это означает, что начиная с номера $n_0 + 1$ последовательность $\{f(x_n)\}$ становится стационарной: $f(x_n) = f(x_0)$ при $n > n_0$. Поэтому существует предел $\lim\limits_{n \to \infty} f(x_n) = f(x_0)$, что, в силу произвольного выбора последовательности $x_n \in X$, $n = 1, 2, \dots$, $\lim\limits_{n \to \infty} x_n = x_0$, означает выполнение условия (5.18), т. е. непрерывность функции f в точке x_0. □

Непрерывность функции в изолированной точке означает, что в математике *дискретное является частным случа-*

ем непрерывного. Например, функция, определенная лишь на одной точке, или на двух точках, или, более общо, на любом конечном множестве точек, является непрерывной. Непрерывной является и элементарная функция $f(x) = 1 + \sqrt{\lg \cos 2\pi x}$ (см. п. 5.2), определенная только для целочисленных значений аргумента x, т. е. для $x = 0, \pm 1, \pm 2, \ldots$, в которых она равна 1 (в остальных точках выражение под знаком корня отрицательно и поэтому функция не определена). График этой функции состоит из отдельных изолированных точек (0, 1), (1, 1), (−1, 1), (2, 1), (−2, 1)... (см. рис. 20).

Из леммы 3 следует, что вопрос о пределе функции в изолированной точке множества ее определения решается совсем просто: он всегда существует и равен $f(x_0)$. Поэтому понятие предела функции (в частности, ее непрерывность) содержательны лишь для предельных точек множества определения функции.

УПРАЖНЕНИЕ 6. Доказать, что функция $f : X \to \boldsymbol{R}$ непрерывна в предельной точке $x_0 \in X$ множества X тогда и только тогда, когда

$$\lim_{\substack{x \to x_0 \\ x \in X \cap \overset{\circ}{U}(x_0)}} f(x) = f(x_0).$$

Подобно тому, как рассматривался предел функции по какому-либо множеству, принадлежащему ее области определения, можно, в частности, рассмотреть и непрерывность функции по соответствующему множеству.

Определение 8. *Пусть $f : X \to \boldsymbol{R}$ и $x_0 \in E \subset X$. Функция f называется непрерывной в точке x_0 по множеству E, если*

$$\lim_{\substack{x \to x_0 \\ x \in E}} f(x) = f(x_0).$$

Иначе говоря, функция f *называется непрерывной в точке $x_0 \in E$ по множеству E, если в этой точке непрерывно сужение f_E этой функции на множестве E*:

$$\lim_{x \to x_0} f_E(x) = f_E(x_0).$$

Например, функция Дирихле f (см. пример 4 в п. 5.4) непрерывна в точке $x_0 = 0$ по множеству $\boldsymbol{Q}$ всех рациональных чисел, так как

$$\lim_{\substack{x \to 0 \\ x \in \boldsymbol{Q}}} f(x) = 1 = f(0),$$

и не является в ней непрерывной по множеству всех действительных чисел, так как предел в точке $x_0 = 0$ по множеству всех действительных чисел у функции Дирихле просто не существует. Очевидно, что функция Дирихле не является непрерывной ни в какой точке числовой оси.

Про функцию $f : X \to \boldsymbol{R}$, непрерывную в точке $x_0 \in X$ в смысле определения 5, можно сказать, что она непрерывна в этой точке по множеству X.

5.6. Условие существования предела функции

Согласно определению предела, функция $f : X \to \boldsymbol{R}$ имеет предел в точке x_0, если, какова бы ни была последовательность $x_n \to x_0$, $x_n \in X$, $n = 1, 2, \dots$, последовательность соответствующих значений функции $\{f(x_n)\}$ имеет предел (конечный или бесконечный), и эти пределы не зависят от выбора указанных последовательностей $\{x_n\}$, т. е. все последовательности $\{f(x_n)\}$ имеют, и при том один и тот же, предел a: $\lim\limits_{n \to \infty} f(x_n) = a$. Это значение a и является пределом функции f в точке x_0.

Покажем, что если предположить несколько меньше, а именно только существование конечного или определенного знака бесконечного предела у каждой рассматриваемой последовательности $\{f(x_n)\}$, то уже из одного этого следует, что все эти пределы совпадают и тем самым функция f в этом случае имеет предел в точке x_0.

ЛЕММА 4. *Для того чтобы функция $f : X \to \boldsymbol{R}$ имела конечный или определенного знака бесконечный предел в точке x_0, являющейся точкой прикосновения множества X, необходимо и достаточно, чтобы для любой последовательности $x_n \to x_0$, $x_n \in X$, $n = 1, 2, \dots$, последовательность соот-*

ветствующих значений функции $\{f(x_n)\}$ *имела конечный или определенного знака бесконечный предел.*

СЛЕДСТВИЕ. *Для того чтобы функция* $f: X \to \boldsymbol{R}$ *имела конечный предел в точке* x_0, *являющейся точкой прикосновения множества* X, *необходимо и достаточно, чтобы для любой последовательности* $x_n \to x_0$, $x_n \in X$, $n = 1, 2, \dots$, *последовательность соответствующих значений функции* $\{f(x_n)\}$ *была сходящейся.*

Доказательство леммы. Необходимость сформулированного в условиях леммы условия для существования предела функции f при $x \to x_0$ содержится в самом определении этого понятия (см. определение 1 в п. 5.4), в котором утверждается существование предела $\lim\limits_{n \to \infty} f(x_n)$ для всех рассматриваемых последовательностей $\{x_n\}$.

Докажем достаточность указанного в лемме условия для существования предела функции. Пусть $f: X \to \boldsymbol{R}$ и для любой последовательности $x_n \to x_0$, $x_n \in X$, $n = 1, 2, \dots$, последовательность $\{f(x_n)\}$ имеет предел (конечный или определенного знака бесконечный).

Рассмотрим какие-либо две последовательности $x'_n \to x_0$ и $x''_n \to x_0$, $x'_n \in X$, $x''_n \in X$, $n = 1, 2, \dots$. Тогда последовательность

$$x_m = \begin{cases} x'_n, \text{ если } m = 2n - 1, \\ x''_n, \text{ если } m = 2n, \end{cases} \quad n = 1, 2, \dots,$$

также имеет своим пределом (конечным или бесконечным) точку x_0 и $x_m \in X$, $m = 1, 2, \dots$. Поэтому, согласно сделанному предположению, существуют пределы $\lim\limits_{n \to \infty} f(x'_n)$, $\lim\limits_{n \to \infty} f(x''_n)$ и $\lim\limits_{m \to \infty} f(x_m)$, причем последовательности $\{f(x'_n)\}$ и $\{f(x''_n)\}$ являются подпоследовательностями последовательности $\{f(x_m)\}$, ибо последовательности $\{x'_n\}$ и $\{x''_n\}$ являются подпоследовательностями последовательности $\{x_m\}$.

Вспомним теперь, что если у некоторой последовательности имеется предел (конечный или бесконечный), то лю-

бая ее подпоследовательность имеет тот же предел (см. п. 4.3). Поэтому

$$\lim_{n \to \infty} f(x_n') = \lim_{m \to \infty} f(x_m), \quad \lim_{n \to \infty} f(x_n'') = \lim_{m \to \infty} f(x_m),$$

откуда

$$\lim_{n \to \infty} f(x_n') = \lim_{n \to \infty} f(x_n'').$$

Таким образом, пределы последовательностей $\{f(x_n)\}$, где $x_n \to x_0$, $x_n \in X$, $n = 1, 2, \ldots$, не зависят от выбора последовательностей $\{x_n\}$. Обозначая общее значение пределов последовательностей $\{f(x_n)\}$ через a будем иметь, согласно определению 1 п. 5.4, что $\lim\limits_{x \to x_0} f(x) = a$.

Утверждение следствия непосредственно следует из леммы 4 (напомним, что термин «сходящаяся последовательность» употребляется только для последовательностей, имеющих конечный предел). □

5.7. Второе определение предела функции

Существует другое определение предела функции, не использующее понятие предела последовательности, а формулируемое в терминах окрестностей и называемое определением предела функции по Коши. Это определение равносильно определению 1 в п. 5.4.

Определение 9. *Точка a называется пределом функции $f: X \to \boldsymbol{R}$ при $x \to x_0$ (или, что то же, в точке x_0), если для любой окрестности $U(a)$ точки a существует такая окрестность $U(x_0)$ точки x_0, что*

$$f(X \cap U(x_0)) \subset U(a). \tag{5.20}$$

Предел функции по Коши также будем обозначать через $\lim\limits_{x \to x_0} f(x)$. Это естественно, так как понятие предела функции по Коши, как это было отмечено выше и будет вскоре доказано, эквивалентно ранее данному в п. 5.4 определению предела функции.

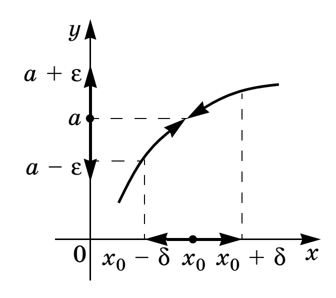


Рис. 23

Рисунок 23 иллюстрирует определение 9 в случае, когда x_0 и a — действительные числа, а множество X является проколотой окрестностью точки x_0.

Используя логические символы, определение 9 можно записать в следующем виде:

$$\lim_{x \to x_0} f(x) = a \overset{\text{def}}{\Leftrightarrow}$$

$$\overset{\text{def}}{\Leftrightarrow} \forall\, U(a) \,\exists\, U(x_0)\colon f(X \cap U(x_0)) \subset U(a).$$

Расшифровывая подробнее включение (5.20), определение 9 можно сформулировать следующим образом.

Точка а называется пределом функции $f: X \to \boldsymbol{R}$ *при* $x \to x_0$, *если для любой окрестности* $U(a)$ *точки а существует такая окрестность* $U(x_0)$ *точки* x_0, *что для любой точки*

$$x \in X \cap U(x_0) \tag{5.21}$$

выполняется включение

$$f(x) \in U(a). \tag{5.22}$$

Используя логические символы, это определение можно записать следующим образом:

$$\lim_{x \to x_0} f(x) = a \overset{\text{def}}{\Leftrightarrow}$$

$$\overset{\text{def}}{\Leftrightarrow} \forall\, U(a) \,\exists\, U(x_0) \,\forall\, x \in X \cap U(x_0)\colon f(x) \in U(a). \tag{5.23}$$

Вспоминая определения окрестностей конечных и бесконечно удаленных точек, определение 9 можно в каждом конкретном случае перефразировать в терминах неравенств. Сформулируем сначала в таком виде определение конечного предела в конечной точке.

Число а называется пределом функции f в точке $x_0 \in \boldsymbol{R}$, *если для любого* $\varepsilon > 0$ *существует такое* $\delta > 0$[1], *что для всех x, удовлетворяющих условиям* $|x - x_0| < \delta$, $x \in X$, *выполняется неравенство* $|f(x) - a| < \varepsilon$.

---

[1] Иногда пишут $\delta = \delta(\varepsilon) > 0$, чтобы подчеркнуть, что выбор δ зависит от ε.

В логических символах это определение выглядит следующим образом:

$$\lim_{x \to x_0} f(x) = a \in \boldsymbol{R} \overset{\text{def}}{\Leftrightarrow}$$
$$\overset{\text{def}}{\Leftrightarrow} \forall\, \varepsilon > 0\ \exists\, \delta > 0\ \forall\, x \in X,\ |x - x_0| < \delta\colon |f(x) - a| < \varepsilon.$$

В частности, если функция f непрерывна в точке $x_0 \in X \subset \boldsymbol{R}$ и $a = f(x_0)$ (в этом случае x_0 и a являются числами), то определение непрерывности в символической записи имеет вид

$$\lim_{x \to x_0} f(x) = f(x_0) \Leftrightarrow$$
$$\Leftrightarrow \forall\, \varepsilon > 0\ \exists\, \delta > 0\ \forall\, x \in X,\ |x - x_0| < \delta\colon |f(x) - f(x_0)| < \varepsilon.$$

Приведем пример определения бесконечных пределов в терминах неравенств. Например, запись $\lim\limits_{x \to +\infty} f(x) = -\infty$ для функции $f: X \to \boldsymbol{R}$ означает, что для любого $\varepsilon > 0$ существует такое $\delta > 0$, что для всех x, удовлетворяющих условию $x > \delta$, $x \in X$, выполняется неравенство $f(x) < -\varepsilon$.

С помощью логических символов это определение записывается следующим образом:

$$\lim_{x \to x_0} f(x) = -\infty \overset{\text{def}}{\Leftrightarrow} \forall\, \varepsilon > 0\ \exists\, \delta > 0\ \forall\, x \in X,\ x > \delta\colon f(x) < -\varepsilon.$$

Из сказанного следует, что могут встречаться различные сочетания перехода к предельным значениям (как конечным, так и бесконечным) зависимых и независимых переменных. Формулировка определения предела функции для каждого отдельного случая в терминах неравенств, или, как еще иногда говорят, на «языке ε и δ», хотя часто и удобнее в конкретных ситуациях, хуже приспособлена к решению общих вопросов, чем определение предела функции в терминах окрестностей, так как требует проведения специальных доказательств отдельно для каждого случая в соответствии с данной формулировкой определения. Поэтому удобнее использовать определение 9, охватывающее все конкретные случаи.

Перейдем теперь к сравнению определений предела функции по Гейне (определение 1) и по Коши (определение 9).

ТЕОРЕМА 1. *Определения* 1 *и* 9 *предела функции в точке прикосновения множества определения функции эквивалентны.*

Доказательство. Докажем сначала, что если функция имеет в некоторой точке предел в смысле определения 1, то она имеет тот же самый предел в этой точке и в смысле определения 9. Пусть $f : X \to \boldsymbol{R}$, x_0 — точка прикосновения множества X и $\lim\limits_{x \to x_0} f(x) = a$ в смысле определения 1. Покажем, что тогда выполняется и условие в правой части формулы (5.23).

Допустим, что это не так, т. е. что

$$\exists\, U(a)\ \forall\, U(x_0)\ \exists\, x \in X \cap U(x_0)\colon f(x) \notin U(a), \tag{5.24}$$

или, иначе говоря, найдется такая окрестность $U(a)$ точки a, что в любой окрестности $U(x_0)$ точки x_0 существует точка $x \in X$, значение функции $f(x)$ в которой не принадлежит к окрестности $U(a)$. В частности, указанные точки x найдутся в каждой окрестности $U\left(x_0, \frac{1}{n}\right)$. Обозначим их через x_n, т. е.

$$x_n \in X \cap U\left(x_0, \frac{1}{n}\right), \tag{5.25}$$

$$f(x_n) \notin U(a),\ n = 1, 2, \ldots . \tag{5.26}$$

Из условия (5.25) следует, что

$$\lim_{n \to \infty} x_n = x_0 \tag{5.27}$$

(см. замечание 1 в п. 4.2). Поскольку $\lim\limits_{x \to x_0} f(x) = a$ в смысле определения 1, то для любой последовательности $x_n \to x_0$, $n = 1, 2, \ldots$, имеет место равенство $\lim\limits_{n \to \infty} f(x_n) = a$. Согласно определению предела последовательности, это означает, что для любой окрестности $U(a)$, в частности и для выбранной выше, существует такой номер n_0, что для всех номеров $n > n_0$ имеет место

$$f(x_n) \in U(a). \tag{5.28}$$

Это противоречит условию (5.26). Полученное противоречие доказывает сделанное утверждение. □

Теперь докажем, что если функция имеет в некоторой точке предел в смысле определения 9, то она имеет в этой точке тот же самый предел и в смысле определения 1. Пусть

$\lim\limits_{x \to x_0} f(x) = a$ в смысле определения 9 предела функции, $f : X \to \boldsymbol{R}$, x_0 — точка прикосновения множества X, и пусть

$$x_n \to x_0,\ x_n \in X,\ n = 1, 2, \dots . \tag{5.29}$$

Покажем, что тогда

$$\lim_{n \to \infty} f(x_n) = a, \tag{5.30}$$

т. е. точка a является и пределом функции f в смысле определения 1.

Зададим произвольную окрестность $U(a)$ точки a и выберем для нее окрестность $U(x_0)$ точки x_0, удовлетворяющую условиям (5.21)—(5.22). Для этой окрестности $U(x_0)$, в силу условия (5.29), найдется такой номер n_0, что для всех номеров $n > n_0$ будет выполняться включение $x_n \in X \cap U(x_0)$. Но тогда, в силу (5.22), имеем $f(x_n) \in U(a)$. Это и означает выполнение условия (5.30). □

Предел функции, как это было отмечено в п. 5.4, является локальным свойством функции в том смысле, что его существование для функции в данной точке (а если он существует, то и его значение) не зависит от сужения функции на пересечение любой окрестности точки x_0 с множеством определения заданной функции. Это хорошо также видно и из определения 9: если задать произвольную окрестность $U_0(x_0)$ точки x_0 и добавить в определение 9 дополнительное условие, состоящее в том, что все окрестности $U(x_0)$, существование которых там утверждается, должны кроме всего прочего содержаться в окрестности $U_0(x_0)$:

$$U(x_0) \subset U_0(x_0),$$

то получится определение, равносильное исходному. Действительно, если условие (5.23) выполняется для некоторой окрестности $U(x_0)$ точки x_0, то оно выполняется и для любой содержащейся в ней окрестности этой точки.

В заключение отметим, что под пределом функции в данной точке обычно понимается конечный предел, если не оговорено что-либо другое, а через $\lim\limits_{x \to x_0} f(x)$, когда ничего не сказано о множестве, по которому берется предел, обознача-

ется конечный или бесконечный предел в точке x_0 по всему множеству определения функции f.

УПРАЖНЕНИЕ 7. Доказать, что если $P(x)$ — многочлен степени $n \geqslant 1$, то $\lim\limits_{x \to -\infty} P(x) = \infty$, $\lim\limits_{x \to +\infty} P(x) = \infty$.

Определение предела функции в точке без труда можно обобщить и на функции, у которых как множества их значений, так и множества их определений принадлежат расширенному множеству действительных чисел $\overline{\boldsymbol{R}} = R \cup \{+\infty\} \cup \{-\infty\}$, т. е. на функции вида $f : X \to \overline{\boldsymbol{R}}$, $X \subset \overline{\boldsymbol{R}}$. Читатель в случае необходимости самостоятельно сформулирует соответствующие определения.

5.8. Предел функции по объединению множеств

Докажем еще одно простое, но полезное для дальнейшего свойство пределов функций по множествам.

ЛЕММА 5. *Пусть* $f : X \to \boldsymbol{R}$, $X_1 \subset X$, $X_2 \subset X$, $X_1 \neq \varnothing$, $X_2 \neq \varnothing$, *и точка* x_0 *является точкой прикосновения множеств* X_1 *и* X_2.

Тогда если в точке x_0 *функция* f *имеет равные пределы (конечные или бесконечные) по множествам* X_1 *и* X_2, *то она имеет в этой точке тот же предел и по их объединению.*

Доказательство. Если

$$\lim_{\substack{x \to x_0 \\ x \in X_1}} f(x) = \lim_{\substack{x \to x_0 \\ x \in X_2}} f(x) = a, \qquad (5.31)$$

то для любой окрестности $U(a)$ точки a существует такая окрестность $U(x_0)$ точки x_0, что образы ее пересечений $X_1 \cap U(x_0)$ и $X_2 \cap U(x_0)$ с множествами X_1 и X_2 содержатся в окрестности $U(a)$, а тогда и образ их объединения $(X_1 \cup X_2) \cap U(x_0)$ также содержится в окрестности $U(a)$. Это и означает, что

$$\lim_{\substack{x \to x_0 \\ x \in X_1 \cup X_2}} f(x) = a. \ \square$$

Пример. Если $\{x_n\}$ — такая последовательность, что $\lim\limits_{k \to \infty} x_{2k} = \lim\limits_{k \to \infty} x_{2k-1} = a$ и a — либо действительное чис-

ло, либо одна из бесконечностей ∞, $+\infty$ или $-\infty$, то и $\lim\limits_{n \to \infty} x_n = a$. Это сразу следует из леммы 5, если рассмотреть функцию $f(n) = x_n$, $n \in \boldsymbol{N}$, и положить $X_1 = \{2k\}$, $X_2 = \{2k - 1\}$, $k = 1, 2, \ldots$.

5.9. Односторонние пределы и односторонняя непрерывность

При изучении функций иногда оказывается полезным рассмотреть пределы их сужений на множествах, соответствующих частному случаю, когда эти множества являются частями множеств определения данных функций, лежащими по одну сторону от точки, в которой рассматривается предел. Такие пределы называются *односторонними пределами*. Это понятие содержательно лишь тогда, когда действительно существуют указанные множества, как с одной, так и с другой стороны от точки x_0, в которой рассматривается предел. В том случае, если точка x_0 является одной из бесконечностей ∞, $+\infty$ или $-\infty$, это заведомо невозможно. Поэтому в настоящем пункте в дальнейшем будем всегда предполагать, что x_0 — действительное число: $x_0 \in \boldsymbol{R}$. Введем для упрощения записи некоторые обозначения.

Для всякого множества $X \subset \boldsymbol{R}$ и для точки $x_0 \in \boldsymbol{R}$ положим

$$X_{\leqslant}(x_0) \overset{\text{def}}{=} \{x : x \in X,\ x \leqslant x_0\},$$

$$X_{\geqslant}(x_0) \overset{\text{def}}{=} \{x : x \in X,\ x \geqslant x_0\}.$$

Иначе говоря, множество $X_{\geqslant}(x_0)$, соответственно $X_{\leqslant}(x_0)$, является пересечением множества X с замкнутым лучом числовой оси, вершиной которого является точка x_0 и который направлен в (отрицательном) положительном направлении.

Определение 10. *Пусть* $f : X \to \boldsymbol{R}$ и $x_0 \in \boldsymbol{R}$. *Точка* a *называется пределом функции* f *слева, соответственно справа, при* $x \to x_0$, *если она является пределом при* $x \to x_0$ *функции* f *по множеству* $X_{\leqslant}(x_0)$, *т. е.*

$$\lim_{\substack{x \to x_0 \\ x \in X_{\leqslant}(x_0)}} f(x) = a,$$

соответственно по множеству $X_{\geqslant}(x_0)$, *т. е.*

$$\lim_{\substack{x \to x_0 \\ x \in X_{\geqslant}(x_0)}} f(x) = a.$$

Для пределов слева и справа функции f по множеству $X \backslash \{x_0\}$ имеются специальные обозначения: предел слева обозначается $f(x_0 - 0)$ или $\lim\limits_{x \to x_0 - 0} f(x)$, а предел справа — $f(x_0 + 0)$ или $\lim\limits_{x \to x_0 + 0} f(x)$. Таким образом,

$$f(x_0 - 0) = \lim_{x \to x_0 - 0} f(x) \overset{\text{def}}{=} \lim_{\substack{x \to x_0 \\ x \in X_{\leqslant}(x_0) \backslash \{x_0\}}} f(x), \qquad (5.32)$$

$$f(x_0 + 0) = \lim_{x \to x_0 + 0} f(x) = \lim_{\substack{x \to x_0 \\ x \in X_{\geqslant}(x_0) \backslash \{x_0\}}} f(x). \qquad (5.33)$$

Если $x_0 = 0$, то вместо $0 + 0$ (вместо $0 - 0$) в случае пределов функций, как и в случае пределов последовательностей (см. п. 4.1), пишут просто $+0$ (соответственно -0).

Пределы слева и справа называются *односторонними пределами*. Если же точка x_0 является верхней гранью для множества $X_{\leqslant}(x_0) \backslash \{x_0\}$ и нижней гранью для $X_{\geqslant}(x_0) \backslash \{x_0\}$ (X — множество определения функции f):

$$x_0 = \sup\,(X_{\leqslant}(x_0) \backslash \{x_0\}) = \inf\,(X_{\geqslant}(x_0) \backslash \{x_0\}),$$

то обычный предел функции f при $x \to x_0$ называется также и ее *двусторонним пределом*.

При определении односторонних пределов удобно использовать понятие «односторонних окрестностей» точки: пересечения $U(x_0) \cap \{x : x \geqslant 0\}$ и $U(x_0) \cap \{x : x \leqslant 0\}$, где $U(x_0)$ — некоторая окрестность точки x_0, называются ее односторонними окрестностями (соответственно правосторонней и левосторонней). Ясно, что определение одностороннего предела функции получится из определения предела функции, если в нем окрестность точки x_0 заменить на ее соответствующую одностороннюю окрестность. Отметим, что пересечения $\overset{\circ}{U}(x_0) \cap \{x : x \geqslant x_0\}$ и $\overset{\circ}{U}(x_0) \cap \{x_0 : x \leqslant x_0\}$, где $\overset{\circ}{U}(x_0)$ — некоторая проколотая окрестность точки x_0, называются ее

проколотыми односторонними окрестностями (соответственно правосторонней и левосторонней).

В качестве примера односторонних пределов рассмотрим функцию $y = \operatorname{sign} x$ (см. пример в п. 5.2 и рис. 19). Пусть $x_n > 0$, $x'_n < 0$, $n = 1, 2, \ldots$, и $\lim\limits_{n \to \infty} x_n = \lim\limits_{n \to \infty} x'_n = 0$. Тогда

$$\lim_{n \to \infty} \operatorname{sign} x_n = \lim_{n \to \infty} 1 = 1,$$

$$\lim_{n \to \infty} \operatorname{sign} x'_n = \lim_{n \to \infty} (-1) = -1.$$

Это означает, что $\lim\limits_{x \to +0} \operatorname{sign} x = 1$ и $\lim\limits_{x \to -0} \operatorname{sign} x = -1$.

Понятие предела слева (справа) при $x \to x_0$, как и вообще понятие предела в точке, содержательно только тогда, когда точка x_0 является точкой прикосновения множества, по которому берется предел, в данном случае множества $X_{\leqslant}(x_0)$ (множества $X_{\geqslant}(x_0)$). Каждое из этих множеств лежит по одну сторону от точки x_0, поэтому она является их точкой прикосновения тогда и только тогда, когда $x_0 = \sup X_{\leqslant}(x_0)$ и, соответственно, $x_0 = \inf X_{\geqslant}(x_0)$.

ТЕОРЕМА 2. *Функция $f: X \to \boldsymbol{R}$ имеет предел в точке $x_0 = \sup X_{\leqslant}(x_0) = \inf X_{\geqslant}(x_0)$, $X_{\leqslant}(x_0) \neq \varnothing$, $X_{\geqslant}(x_0) \neq \varnothing$, в том и только том случае, когда в этой точке у функции f существуют пределы как слева, так и справа и они равны. В этом случае их общее значение и является пределом функции в точке x_0.*

СЛЕДСТВИЕ. *Для того чтобы у функции $f: X \to \boldsymbol{R}$ существовал в точке x_0 двусторонний предел по множеству $X \backslash \{x_0\}$, необходимо и достаточно, чтобы существовали и были равны между собой односторонние пределы $f(x_0 - 0)$ и $f(x_0 + 0)$.*

Доказательство теоремы. В самом деле, пусть $f: X \to \boldsymbol{R}$, $x_0 = \sup X_{\leqslant}(x_0) = \inf X_{\geqslant}(x_0)$ и $\lim\limits_{x \to x_0} f(x) = a$, т. е. у функции f существует предел (по множеству X) в точке x_0. Но тогда в этой точке тот же предел существует и у ее суже-

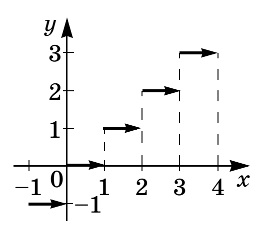


Рис. 24

ния по любому множеству (см. лемму 1 в п. 5.4), в частности и по множествам $X_{\leqslant}(x_0)$ и $X_{\geqslant}(x_0)$, т. е. существуют оба односторонних предела при $x \to x_0$ и они равны a:

$$\lim_{\substack{x \to x_0 \\ x \in X_{\leqslant}(x_0)}} f(x) = \lim_{\substack{x \to x_0 \\ x \in X_{\geqslant}(x_0)}} f(x) = a. \quad (5.34)$$

Пусть, наоборот, в точке $x_0 = \sup X_{\leqslant}(x_0) = \inf X_{\geqslant}(x_0)$ выполняется условие (5.34). Тогда так как $X = X_{\leqslant}(x_0) \cup X_{\geqslant}(x_0)$, то согласно лемме 5, существует предел $\lim\limits_{\substack{x \to x_0 \\ x \in X}} f(x) = a$.

Для того чтобы убедиться в справедливости следствия, достаточно применить теорему 2 к сужению функции $f : X \to \boldsymbol{R}$ на множество $X \backslash \{x_0\}$. □

Если один из односторонних пределов функции в некоторой точке совпадает со значением функции в этой точке, то такая функция называется *односторонне непрерывной в рассматриваемой точке*. Сформулируем это определение более подробно.

Определение 11. *Функция* $f : X \to \boldsymbol{R}$ *называется непрерывной слева, соответственно справа, в точке* $x_0 \in X$, *если* $\lim\limits_{\substack{x \to x_0 \\ x \in X_{\leqslant}(x_0)}} f(x) = f(x_0)$, *соответственно если* $\lim\limits_{\substack{x \to x_0 \\ x \in X_{\geqslant}(x_0)}} f(x) = f(x_0)$.

Пример. Рассмотрим функцию, определенную на всей числовой оси и равную для каждого действительного числа x наибольшему целому числу, меньшему или равному x. Эта функция имеет специальное обозначение $y = [x]$, читается: «y является целой частью числа x» или «y равно entier x»[1].

Таким образом,

$$[x] = n \in \boldsymbol{Z}, \;\; n \leqslant x, \;\; \text{а} \;\; n + 1 > x.$$

График функции $[x]$ изображен на рисунке 24. Функция $[x]$ в целочисленных точках $x = n$, $n = 0, \pm 1, \pm 2, \ldots$, числовой прямой непрерывна справа и разрывна слева. Во всех же других точках она непрерывна как справа, так и слева. Таким образом, в частности, функция $[x]$ непрерывна справа во всех точках числовой оси.

[1] entier (франц.) — целый.

З а м е ч а н и е. Если для функции $f: X \to \boldsymbol{R}$ существует конечный предел $\lim\limits_{x \to x_0} f(x) = a$, причем для всех $x \in X$ выполняется неравенство $f(x) < a$ (неравенство $f(x) > a$), то пишут $\lim\limits_{x \to x_0} f(x) = a - 0$ (соответственно $\lim\limits_{x \to x_0} f(x) = a + 0$). При этом если $a = 0$, то также вместо $0 + 0$ и $0 - 0$ пишут просто $+0$ и -0.

5.10. Свойства пределов функций

Все функции, рассматриваемые в этом пункте, определены на некотором множестве $X \subset \boldsymbol{R}$, и все их пределы берутся по множеству X в некоторой точке x_0, которая является точкой прикосновения (конечной или бесконечно удаленной) множества X. Напомним, что если x_0 — действительное число: $x_0 \in \boldsymbol{R}$, то x_0 — обычная точка прикосновения множества X. Если $x_0 = \infty$, то множество X не ограничено, а если $x_0 = +\infty$, или $x_0 = -\infty$, то множество X не ограничено сверху, соответственно снизу.

1^0. *Если функция $f: X \to \boldsymbol{R}$ имеет в точке x_0 конечный предел, то существует такая окрестность $U(x_0)$ точки x_0, что функция f ограничена на пересечении $U(x_0) \cap X$ этой окрестности с множеством определения X функции f.*

СЛЕДСТВИЕ. *Функция $f: X \to \boldsymbol{R}$, непрерывная в точке $x_0 \in X$, ограничена на пересечении некоторой окрестности этой точки с множеством X.*

Д о к а з а т е л ь с т в о. Пусть $\lim\limits_{x \to x_0} f(x) = a$ — конечный предел, тогда, согласно определению 9 из п. 5.7, для любого $\varepsilon > 0$, в частности для $\varepsilon = 1$, существует такая окрестность $U(x_0)$ точки x_0, что $f(X \cap U(x_0)) \subset U(a, 1)$, т. е. для всех $x \in X \cap U(x_0)$ выполняется включение $f(x) \in U(a, 1)$. Иначе говоря, для всех $x \in X \cap U(x_0)$ справедливо неравенство $a - 1 < f(x) < a + 1$,

а это означает ограниченность функции f на пересечении $X \cap U(x_0)$. □

Следствие непосредственно вытекает из доказанного утверждения, так как непрерывность функции в точке является частным случаем существования у нее в точке конечного предела.

2^0. *Если функция* $f : X \to \boldsymbol{R}$ *имеет в точке* x_0 *не равный нулю конечный предел* $\lim\limits_{x \to x_0} f(x) = a \neq 0$, *то для любого числа* c, $0 < c < |a|$, *существует такая окрестность* $U(x_0)$ *точки* x_0, *что для всех точек* x *из области определения* X *функции* f, *принадлежащих окрестности* $U(x_0)$, *т. е. для всех* $x \in U(x_0) \cap X$, *выполняются неравенства* $f(x) > c$, *если* $a > 0$ *и*

$$f(x) < -c, \text{ если } a < 0. \tag{5.36}$$

СЛЕДСТВИЕ **(лемма о сохранении знака).** *Если функция* $f : X \to \boldsymbol{R}$ *непрерывна в точке* $x_0 \in X$ *и* $f(x_0) > 0$ (*соответственно* $f(x_0) < 0$), *то существует такая окрестность* $U(x_0)$ *точки* x_0, *что для всех точек* $x \in X \cap U(x_0)$ *выполняется неравенство* $f(x) > 0$ (*соответственно* $f(x) < 0$).

Доказательство свойства 2^0. Пусть сначала $a > 0$ и $0 < c < a$. Выберем $\varepsilon > 0$ так, чтобы $c < a - \varepsilon$, тогда $c \notin U(a; \varepsilon) = (a - \varepsilon, a + \varepsilon)$. Для окрестности $U(a; \varepsilon)$ существует такая окрестность $U(x_0)$ точки x_0, что для любой точки $x \in U(x_0) \cap X$ имеет место включение $f(x) \in (a - \varepsilon, a + \varepsilon)$ и, следовательно, $f(x) > a - \varepsilon > c$.

Аналогично, если $a < 0$ и $0 < c < -a$, т. е. $a < -c$, то выберем $\varepsilon > 0$ так, чтобы $a + \varepsilon < -c$, тогда $c \notin U(a; \varepsilon) = (a - \varepsilon, a + \varepsilon)$. Для окрестности $U(a; \varepsilon)$ существует такая окрестность $U(x_0)$ точки x_0, что для любой точки $x \in U(x_0) \cap X$ имеет место включение $f(x) \in (a - \varepsilon, a + \varepsilon)$, поэтому и неравенство $f(x) < < a + \varepsilon < -c$.

Следствие непосредственно вытекает из неравенств (5.36). □

Замечание. Если в точке x_0 у функции $f : X \to \boldsymbol{R}$ существует бесконечный предел, равный ∞, $+\infty$ или $-\infty$, то для любого $c > 0$ имеет место утверждение, аналогичное

свойству 2^0. Это непосредственно следует из определения бесконечного предела функции, сформулированного в терминах неравенств. А именно: $\lim\limits_{x \to x_0} f(x) = \infty$ (соответственно $+\infty$ или $-\infty$) означает, что для любого $c > 0$ существует такая окрестность $U(x_0)$ точки x_0, что для всех $x \in X \cap U(x_0)$ выполняется неравенство $|f(x)| > c$ (соответственно неравенство $f(x) > c$ или $f(x) < -c$).

3^0. *Если $f(x) = c$ — постоянная, $x \in X$, то* $\lim\limits_{x \to x_0} f(x) = c$.

4^0. *Если $f(x) \geqslant a$, $x \in X$, и существует конечный или определенного знака бесконечный предел $\lim\limits_{x \to x_0} f(x)$, то*

$$\lim_{x \to x_0} f(x) \geqslant a. \tag{5.37}$$

5^0. *Если $\varphi(x) \leqslant f(x) \leqslant \psi(x)$, $x \in X$, и существуют конечные или определенного знака бесконечные пределы $\lim\limits_{x \to x_0} \varphi(x) = \lim\limits_{x \to x_0} \psi(x) = a$, то*

$$\lim_{x \to x_0} f(x) = a. \tag{5.38}$$

6^0. *Если существуют конечные пределы $\lim\limits_{x \to x_0} f(x)$ и $\lim\limits_{x \to x_0} g(x)$, то существуют и конечные пределы $\lim\limits_{x \to x_0} [f(x) + g(x)]$, $\lim\limits_{x \to x_0} f(x)g(x)$, а если $\lim\limits_{x \to x_0} g(x) \neq 0$, то и предел $\lim\limits_{x \to x_0} \frac{f(x)}{g(x)}$, причем*

$$\lim_{x \to x_0} [f(x) + g(x)] = \lim_{x \to x_0} f(x) + \lim_{x \to x_0} g(x), \tag{5.39}$$

$$\lim_{x \to x_0} f(x)g(x) = \lim_{x \to x_0} f(x) \lim_{x \to x_0} g(x), \tag{5.40}$$

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \frac{\lim\limits_{x \to x_0} f(x)}{\lim\limits_{x \to x_0} g(x)}. \tag{5.41}$$

СЛЕДСТВИЕ 1. *Если существует конечный предел* $\lim\limits_{x \to x_0} f(x)$, *то для любого числа* $x \in \boldsymbol{R}$ *существует и предел* $\lim\limits_{x \to x_0} cf(x)$, *причем*

$$\lim_{x \to x_0} cf(x) = c \lim_{x \to x_0} f(x). \qquad (5.42)$$

В случае $c \neq 0$ *равенство* (5.42) *справедливо и для бесконечных пределов определенного знака.*

СЛЕДСТВИЕ 2. *Если функции* f *и* g *непрерывны в точке* $x_0 \in X$, *то функции* cf (c *— постоянная*), $f + g$, fg, *а если, кроме того,* $g(x_0) \neq 0$, *то и функция* $\frac{f}{g}$ *также непрерывна в точке* x_0.

Заметим, что при предположениях, в которых сформулированы утверждение 6 и следствие 2 из него, частное $\frac{f(x)}{g(x)}$, конечно, может быть не определено на всем исходном множестве X, так как на нем могут существовать точки x, в которых $g(x) = 0$. Однако, согласно свойству 2^0, из условия $\lim\limits_{x \to x_0} g(x) \neq 0$ следует, что существует такая окрестность $U(x_0)$ точки x_0, на пересечении которой с множеством X выполняется неравенство $g(x) \neq 0$ и, следовательно, на этом пересечении уже определено частное $\frac{f(x)}{g(x)}$. В формуле (5.41) под пределом подразумевается предел сужения функции $\frac{f(x)}{g(x)}$ на множество $U(x_0) \cap X$. В силу того что предел функции в точке является локальным свойством (см. п. 5.4 и п. 5.7), этот предел не зависит от выбора указанной окрестности $U(x_0)$.

Свойства 3^0—6^0 могут быть доказаны одинаковым методом, основанным на соответствующих свойствах пределов последовательностей (см. п. 4.9).

Докажем, например, формулу (5.40). Пусть $a = \lim\limits_{x \to x_0} f(x)$, $b = \lim\limits_{x \to x_0} g(x)$. Тогда, согласно определению 1 предела функ-

ции (см. п. 5.4), для любой последовательности $x_n \in X$, $n = 1, 2, \ldots$, $\lim\limits_{n \to \infty} x_n = x_0$, справедливы равенства

$$a = \lim_{n \to \infty} f(x_n), \ b = \lim_{n \to \infty} g(x_n).$$

Поэтому, вспомнив, что предел произведения сходящихся последовательностей существует и равен произведению их пределов (см. п. 4.9), получим, что существует предел

$$\lim_{n \to \infty} f(x_n)g(x_n) = ab. \qquad (5.43)$$

Этот предел не зависит от выбора указанной последовательности $\{x_n\}$: он всегда равен ab, поэтому, согласно тому же определению 1, доказанное равенство (5.43) и означает, что

$$\lim_{x \to x_0} f(x)g(x) = ab = \lim_{x \to x_0} f(x) \lim_{x \to x_0} g(x). \ \square$$

Следствие 1 (в силу свойства 3^0) представляет собой частный случай формулы (5.40). Следствие 2 непосредственно вытекает из свойства 6^0, поскольку непрерывность функции в точке означает существование у нее в этой точке конечного предела, равного значению функции в этой точке. Например,

$$\lim_{x \to x_0} f(x)g(x) = \lim_{x \to x_0} f(x) \lim_{x \to x_0} g(x) = f(x_0)g(x_0), \qquad (5.44)$$

так как пределы $\lim\limits_{x \to x_0} f(x)$ и $\lim\limits_{x \to x_0} g(x)$, в силу непрерывности функций f и g в точке x_0, равны соответственно $f(x_0)$ и $g(x_0)$. Выполнение равенства (5.44) и означает непрерывность произведения fg в точке x_0. $\square$

Отметим, что приведенное доказательство следствия 2 можно было бы и не проводить, так как непрерывность функции в точке означает (см. п. 5.5), что эта точка принадлежит множеству задания функции и что у функции в этой точке существует предел по указанному множеству. Поскольку функции f и g заданы в точке x_0, очевидно, и функции $f(x) + g(x)$, $f(x)g(x)$, а при $g(x_0) \neq 0$ и функция $\frac{f(x)}{g(x)}$ заданы в этой точке. В силу свойства 6^0 у перечисленных функ-

ций существуют пределы в точке x_0, принадлежащей в данном случае их множеству задания, что и означает их непрерывность в этой точке. Иначе говоря, утверждение следствия 2 утверждения 6^0 является просто частным случаем этого утверждения, когда точка, в которой рассматривается предел, принадлежит области задания функций.

5.11. Бесконечно малые и бесконечно большие функции

Все рассматриваемые в этом пункте функции будем предполагать определенными на множестве $X \subset \boldsymbol{R}$ и рассматривать их конечные и бесконечные пределы при стремлении аргумента к конечной или к бесконечно удаленной точке x_0.

Определение 12. *Функция* $\alpha\colon X \to \boldsymbol{R}$ *называется бесконечно малой при* $x \to x_0$, *если*

$$\lim_{x \to x_0} \alpha(x) = 0. \tag{5.45}$$

Бесконечно малые функции играют особую роль среди всех функций, имеющих предел, связанную, в частности, с тем, что общее понятие конечного предела может быть сведено к понятию бесконечно малой. Сформулируем это утверждение в виде леммы.

ЛЕММА 6. *Конечный предел* $\lim\limits_{x \to x_0} f(x)$ *существует и равен* a *тогда и только тогда, когда* $f(x) = a + \alpha(x)$, $x \in X$, *где* $\alpha = \alpha(x)$ — *бесконечно малая при* $x \to x_0$.

Доказательство. Если $\lim\limits_{x \to x_0} f(x) = a$, то, положив $\alpha(x) = f(x) - a$, $x \in X$, получим, что

$$\lim_{x \to x_0} \alpha(x) = \lim_{x \to x_0} f(x) - a = a - a = 0.$$

Наоборот, если $f(x) = a + \alpha(x)$, $x \in X$ и $\lim\limits_{x \to x_0} \alpha(x) = 0$, то

$$\lim_{x \to x_0} f(x) = a + \lim_{x \to x_0} \alpha(x) = a. \ \square$$

ТЕОРЕМА 3. *Сумма и произведение конечного числа бесконечно малых при* $x \to x_0$, *а также и произведение бесконеч-*

но малой при $x \to x_0$ на ограниченную на X функцию являются бесконечно малыми при $x \to x_0$.

Доказательство. То, что сумма и произведение конечного числа бесконечно малых являются бесконечно малыми, непосредственно следует из свойства суммы и произведения пределов функций (см. свойство 6 в п. 5.10) в том частном случае, когда эти пределы равны нулю.

Докажем последнее утверждение теоремы. Пусть $\lim\limits_{x \to x_0} \alpha(x) = 0$ и $f(x)$ — ограниченная функция, т. е. существует такая постоянная $b > 0$, что для всех $x \in X$ выполняется неравенство $|f(x)| \leqslant b$. Если $x_n \in X$, $n = 1, 2, \ldots$, — такая последовательность, что $\lim\limits_{n \to \infty} x_n = x_0$, то, согласно определению 1 предела функции (см. п. 5.4), имеем $\lim\limits_{n \to \infty} \alpha(x_n) = 0$. Для всех $n = 1, 2, \ldots$ выполняется неравенство $|f(x_n)| \leqslant b$, т. е. последовательность $\{f(x_n)\}$ ограничена. Но произведение бесконечно малой последовательности, в данном случае последовательности $\{\alpha(x_n)\}$, на ограниченную последовательность, в данном случае на $\{f(x_n)\}$, является бесконечно малой последовательностью (см. свойство 2^0 в п. 4.8), поэтому $\lim\limits_{n \to \infty} f(x_n)\, \alpha(x_n) = 0$. Так как это верно для любой указанной последовательности $\{x_n\}$, то согласно тому же определению предела функции получим $\lim\limits_{x \to x_0} f(x)\, \alpha(x) = 0$, а это и означает, что функция $f(x)\, \alpha(x)$ является бесконечно малой при $x \to x_0$. □

Наряду с бесконечно малыми в анализе часто встречаются бесконечно большие функции. Определим их.

Определение 13. *Функция $f : X \to \boldsymbol{R}$ называется бесконечно большой при $x \to x_0$, если*

$$\lim_{x \to x_0} f(x) = \infty. \qquad (5.46)$$

Между бесконечно большими и бесконечно малыми существует тесная связь. Именно величина, обратная беско-

нечно большой, является бесконечно малой, и наоборот. Более точно, справедливо, например, следующее утверждение.

ЛЕММА 7. *Если функция* $f: X \to \overline{\boldsymbol{R}}$ *бесконечно большая при* $x \to x_0$, *то функция* $\frac{1}{f}$ *является бесконечно малой при* $x \to x_0$.

Доказательство. Пусть произвольно фиксировано $\varepsilon > 0$. Тогда, согласно условию $\lim\limits_{x \to x_0} f(x) = \infty$, существует такая окрестность $U(x_0)$ точки x_0, что для всех точек $x \in U(x_0) \cap X$ выполняется неравенство $|f(x)| > \frac{1}{\varepsilon}$, следовательно, и неравенство $\left|\frac{1}{f(x)}\right| < \varepsilon$. А это и означает, что $\lim\limits_{x \to x_0} \frac{1}{f(x)} = 0$, т. е. что функция является бесконечно малой. □

Замечание 1. Как всегда, когда речь идет о частном функций со знаменателем, предел которого отличен от нуля, здесь под $\frac{1}{f(x)}$ понимается, вообще говоря, частное от деления 1 на сужение функции f на пересечение $U(x_0) \cap X$ такой окрестности $U(x_0)$ точки x_0 с множеством определения X функции f, что для всех точек $x \in U(x_0) \cap X$ функция f не равна нулю. Существование указанной окрестности $U(x_0)$ следует из свойства 2^0 пределов функций (см. п. 5.10). Впрочем, оно еще раз для случая бесконечно большой функции было получено в процессе доказательства леммы 7: очевидно, что из условия $|f(x)| > \frac{1}{\varepsilon}$, $\varepsilon > 0$, следует, что $f(x) \neq 0$.

Замечание 2. Если, наоборот, $\alpha(x)$ — бесконечно малая при $x \to x_0$ функция, то может случиться, что обратная величина $\frac{1}{\alpha(x)}$ не будет определена на множестве, для которого точка x_0 является точкой прикосновения (например, это заведомо имеет место при $\alpha(x) \equiv 0$ на X), и поэтому понятие предела $\lim\limits_{x \to x_0} \frac{1}{\alpha(x)}$ при $x \in X$ будет бессодержательно. Однако если X_0 — такое подмножество множества X, на котором $\alpha(x) \neq 0$: $X_0 = \{x: x \in X, \alpha(x) \neq 0\}$, и если x_0 является точ-

кой прикосновения множества X_0, то функция $\frac{1}{\alpha(x)}$ определена на X_0 и на этом множестве $\lim\limits_{\substack{x \to x_0 \\ x \in X_0}} \frac{1}{\alpha(x)} = \infty$. Именно в этом смысле говорят, что функция, обратная бесконечно малой, является бесконечно большой.

То обстоятельство, что функция, обратная бесконечно малой, является бесконечно большой, и наоборот, делает естественными следующие символические обозначения, часто употребляющиеся для сокращения записи: для любого числа $a > 0$ пишут:

$$\frac{a}{+0} = +\infty, \quad \frac{a}{-0} = -\infty, \quad \frac{\pm a}{0} = \infty,$$
$$\frac{a}{+\infty} = +0, \quad \frac{a}{-\infty} = -0, \quad \frac{\pm a}{\infty} = 0. \qquad (5.47)$$

З а м е ч а н и е 3. На бесконечно большие функции свойства конечных пределов, связанные с арифметическими действиями над пределами (см. свойство 6^0 в п. 5.10), непосредственно не переносятся. Однако некоторые аналогии имеют место.

Например, если $\lim\limits_{x \to x_0} f(x) = +\infty$, $\lim\limits_{x \to x_0} g(x) = +\infty$, то и $\lim\limits_{x \to x_0} [f(x) + g(x)] = +\infty$. Однако о существовании какого-либо предела $\lim\limits_{x \to x_0} [f(x) - g(x)]$ здесь, вообще говоря, уже ничего утверждать нельзя. Можно показать, что «позитивные» утверждения о бесконечных пределах имеют место в случаях, для которых формулами (5.47) и в п. 3.1 были определены некоторые «арифметические операции» с бесконечностями.

5.12. Различные формы записи непрерывности функции в точке

Условие непрерывности в точке x_0 (см. (5.18))

$$\lim_{x \to x_0} f(x) = f(x_0)$$

функции f, заданной на множестве X, можно понимать как в смысле определения предела функции по Гейне (см. опре-

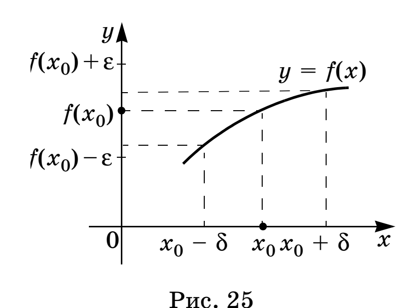


Рис. 25

деление 1 в п. 5.4), так и в смысле определения по Коши (см. определение 9 в п. 5.7). В первом случае это означает, что для любой последовательности

$$x_n \in X,\ n = 1, 2, \ldots,$$

$$\lim_{n \to \infty} x_n = x_0, \qquad (5.48)$$

выполняется условие

$$\lim_{n \to \infty} f(x_n) = f(x_0). \qquad (5.49)$$

Во втором случае это означает, что для любого $\varepsilon > 0$ существует такое $\delta > 0$, что для всех точек x, удовлетворяющих условию

$$|x - x_0| < \delta,\ x \in X, \qquad (5.50)$$

выполняется (рис. 25) неравенство

$$|f(x) - f(x_0)| < \varepsilon. \qquad (5.51)$$

Понятие непрерывности функции, сформулированное в терминах последовательностей (определение (5.48)—(5.49)), отражает собой ситуацию, часто встречающуюся на практике при косвенном вычислении какой-либо величины y, т. е. вычислении ее с помощью измерения некоторого параметра x, от которого эта величина непрерывно зависит, $y = f(x)$. Именно знание непрерывности функции $y = f(x)$ в точке x_0 дает объективную уверенность в том, что чем точнее будут последовательно получаться (в результате экспериментов, измерений или расчетов) значения x_n, $n = 1, 2, \ldots$, приближающие значение x_0, тем точнее будут и соответствующие приближенные значения $y_n = f(x_n)$ величины $y_0 = f(x_0)$.

Определение (5.50)—(5.51) непрерывности функции f в точке x_0 можно еще перефразировать так: функция f непрерывна в точке x_0, если, какова бы ни была заданная степень точности $\varepsilon > 0$ для значений функции f, существует такая степень точности $\delta = \delta(\varepsilon) > 0$ для аргумента, что коль скоро мы выберем значение аргумента x, равное x_0 с точностью до δ, т. е. удовлетворяющее условию (5.50), и возьмем значение функции $f(x)$, то получим значение $f(x_0)$ с заданной сте-

пенью точности, т. е. будет выполнено неравенство (5.51). Это высказывание является, конечно, перефразировкой определения (5.50)—(5.51), разъясняющей интуитивное представление о непрерывной функции.

Так как непрерывность функции в точке является частным случаем существования предела функции, то определение непрерывности функции в точке можно дать в терминах окрестностей, надо лишь к условию (5.20) добавить требование $x_0 \in X$. Таким образом, функция f, определенная на множестве X, непрерывна в точке x_0, если для любой окрестности $U(y_0)$ точки $y_0 = f(x_0)$ существует такая окрестность $U(x_0)$ точки x_0, что выполняется включение

$$f(U(x_0) \cap X) \subset U(y_0), \quad x_0 \in X. \tag{5.52}$$

Наконец, перенося постоянную $f(x_0)$ в равенстве (5.18) в левую часть, внося ее под знак предела и замечая, что обозначение $x \to x_0$ при пределе функции равносильно обозначению $x - x_0 \to 0$ (см. п. 5.4), получим

$$\lim_{x - x_0 \to 0} [f(x) - f(x_0)] = 0. \tag{5.53}$$

Разность $x - x_0$ называется *приращением аргумента* и обозначается через Δx, а разность $f(x) - f(x_0)$ — *приращением функции* $y = f(x)$, соответствующим данному приращению аргумента Δx, и обозначается через Δy. Таким образом,

$$\Delta x = x - x_0,\ \Delta y = f(x_0 + \Delta x) - f(x_0),\ x_0 \in X,\ x \in X. \tag{5.54}$$

В этих обозначениях равенство (5.53) принимает вид

$$\lim_{\Delta x \to 0} \Delta y = 0, \tag{5.55}$$

т. е. непрерывность функции в точке означает, что бесконечно малому приращению аргумента соответствует бесконечно малое приращение функции.

Иногда бывает полезным условие непрерывности функции в точке, основанное на рассмотрении предела функции по проколотой окрестности.

ЛЕММА 3. *Если существует предел*

$$\lim_{\substack{x \to x_0 \\ x \in \mathring{U}(x_0) \cap X}} f(x) = a \qquad (5.56)$$

и функция f определена в точке x_0, то f непрерывна в точке x_0 тогда и только тогда, когда $f(x_0) = a$.

Доказательство. Если функция f непрерывна в точке x_0, т. е. выполняется условие (5.13), то в силу очевидного включения $\mathring{U}(x_0) \cap X \subset X$ и того, что из существования предела по множеству следует существование предела и по любому подмножеству, имеем

$$\lim_{\substack{x \to x_0 \\ x \in \mathring{U}(x_0) \cap X}} f(x) = f(x_0). \qquad (5.57)$$

Из (5.56) и (5.57) следует, что $f(x_0) = a$.

Пусть теперь, наоборот, выполняется условие $f(x_0) = a$ и, следовательно,

$$\lim_{\substack{x \to x_0 \\ x \in \mathring{U}(x_0) \cap X}} f(x) \underset{(5.56)}{=} f(x_0).$$

Отсюда имеем, что для любой окрестности $U(f(x_0))$ точки $f(x_0)$ существует такая окрестность $U(x_0)$ точки x_0, что

$$f(\mathring{U}(x_0) \cap X) \subset U(f(x_0)). \qquad (5.58)$$

Но, очевидно, $f(x_0) \in U(f(x_0))$, поэтому в левой части включения (5.58) можно проколотую окрестность $\mathring{U}(x_0)$ заменить обычной окрестностью $U(x_0)$:

$$f(U(x_0) \cap X) \subset U(f(x_0)).$$

Это и означает, что функция f непрерывна в точке x_0. □

Примеры. 1. Функция $f(x) = c$, где c — постоянная, непрерывна на всей числовой прямой.

В самом деле, для любого $x_0 \in \boldsymbol{R}$ имеет место равенство

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} c = c = f(x_0). \ \square$$

2. Функция $f(x) = \frac{1}{x}$ непрерывна в каждой точке $x_0 \neq 0$.

В самом деле,

$$\Delta y = f(x_0 + \Delta x) - f(x_0) = \frac{1}{x + \Delta x} - \frac{1}{x_0} = -\frac{\Delta x_0}{(x_0 + \Delta x)x_0},$$

откуда при $x_0 \neq 0$ имеем

$$\lim_{\Delta x \to 0} \Delta y = -\lim_{\Delta x \to 0} \frac{\Delta x}{(x_0 + \Delta x)x_0} = -\frac{\lim\limits_{\Delta x \to 0} \Delta x}{\lim\limits_{\Delta x \to 0} (x_0 + \Delta x)x_0} = -\frac{0}{x_0^2} = 0.$$

Это, согласно (5.52), и означает непрерывность функции $f(x) = \frac{1}{x}$ в точке $x_0 \neq 0$. □

3. Функция $f(x) = |\text{sign } x|$ (см. рис. 22) не является непрерывной в точке $x_0 = 0$, так как предел этой функции (по всей числовой оси) в точке $x_0 = 0$ не существует (см. пример 5 в п. 5.4).

УПРАЖНЕНИЯ. 8. Выяснить, с какой степенью точности достаточно задать значения аргумента функции x^3 в данной точке x_0, чтобы получить значение функции с заданной степенью точности $\varepsilon > 0$.

9. Выяснить, является ли функция

$$f(x) = \begin{cases} x \cos \frac{1}{x} & \text{при } x \neq 0, \\ 0 & \text{при } x = 0, \end{cases}$$

непрерывной в точке $x = 0$.

До сих пор в качестве примеров функций, определенных на отрезках и имеющих разрывы, мы рассматривали функции, у которых имелось либо конечное множество точек разрыва (например, функция $y = \text{sign } x$ имеет на отрезке $[-1, 1]$ одну точку разрыва $x = 0$), либо множество точек разрыва совпадало со всем отрезком (функция Дирихле). Приведем теперь пример функции, имеющей бесконечное множество точек разрыва, не совпадающее со всем отрезком, на котором задана функция.

Пример 4. Пусть функция f равна $\frac{1}{n}$, $n \in N$, в каждой не равной нулю рациональной точке $r = \frac{m}{n}$ $(m \neq 0)$ отрезка $[0, 1]$, где $\frac{m}{n}$ — несократимая рациональная дробь, и равна нулю во всех остальных точках отрезка, т. е. в нуле и в иррациональных точках. Эта функция называется *функцией Римана*[1]. Функция f разрывна в каждой рациональной точке $r \neq 0$, так как для любой такой точки существует последовательность иррациональных чисел ξ_n, $n = 1, 2, \ldots$, стремящихся к r, для которых, согласно определению функции

[1] Б. Риман (1826—1866) — немецкий математик.

Римана, $f(\xi_n) = 0$ и, следовательно, $\lim\limits_{n \to \infty} f(\xi_n) = 0 \neq f(r)$, $\lim\limits_{n \to \infty} \xi_n = r$, т. е. функция Римана разрывна во всех рациональных точках $r \neq 0$.

Покажем, что в каждой иррациональной точке и в нуле функция Римана непрерывна. Пусть число ξ либо нуль, либо иррационально. Зададим произвольно $\varepsilon > 0$ и выберем натуральное n так, чтобы $\frac{1}{n} < \varepsilon$. Обозначим через δ расстояние от точки ξ до ближайшего из чисел $\frac{1}{m}, \frac{2}{m}, \ldots, \frac{m-1}{m}, 1$ ($m = 1, 2, \ldots, n$). Очевидно, $\delta > 0$, так как ξ не равно ни одному из указанных чисел и их конечное множество. Пусть $|x - \xi| < \delta$ и x — рациональное число, т. е. выражается несократимой рациональной дробью $x = \frac{p}{q}$; тогда, согласно выбору δ, имеет место неравенство $q > n$, поэтому $|f(x) - f(\xi)| = \frac{1}{q} < \frac{1}{n} < \varepsilon$, ибо $f(\xi) = 0$. Если же $|x - \xi| < \delta$ и x — иррациональное число, то $f(x) = 0$ и $|f(x) - f(\xi)| = 0 < \varepsilon$. Это и означает непрерывность функции Римана в точке ξ. □

5.13. Классификация точек разрыва функции

Определение 14. *Пусть функция f определена в некоторой окрестности точки x_0, кроме, быть может, самой этой точки.*

Точка x_0 называется точкой разрыва функции f, если функция f не определена в точке x_0 или если она определена в этой точке, но не является в ней непрерывной.

УПРАЖНЕНИЕ 10. Сформулировать в «позитивном» смысле определение точки разрыва функции.

Образно говоря, точка x_0 является точкой разрыва функции, если x_0 является значением аргумента, при котором происходит «разрыв графика функции».

Определение 15. *Если x_0 — точка разрыва функции f и существуют конечные односторонние пределы*[1] $f(x_0 - 0) =$

---

[1] Напомним, что односторонние пределы $f(x_0 - 0)$ и $f(x_0 + 0)$ берутся по множеству, не содержащему саму точку x_0.

$= \lim\limits_{x \to x_0 - 0} f(x)$ *и* $f(x_0 + 0) = \lim\limits_{x \to x_0 + 0} f(x)$, *то точка* x_0 *называется точкой разрыва первого рода.*

Величина $f(x_0 + 0) - f(x_0 - 0)$ называется *скачком функции* f в точке x_0. Если скачок функции f в точке разрыва x_0 равен нулю, т. е. $f(x_0 + 0) = f(x_0 - 0)$, то x_0 называется *точкой устранимого разрыва*.

Последний термин оправдан тем, что если в этом случае переопределить или доопределить (если функция f была не определена в точке x_0) функцию f, положив

$$f(x_0) = f(x_0 - 0) = f(x_0 + 0), \tag{5.59}$$

то получим непрерывную в точке x_0 функцию.

Действительно, покажем, что если для функции $f : X \to \boldsymbol{R}$ выполнено условие (5.59), то она непрерывна в точке x_0. Положим $X_1 = X \backslash \{x_0\}$ и $X_2 = \{x_0\}$. В силу теоремы 2 п. 5.9, из равенства $f(x_0 + 0) = f(x_0 - 0)$ следует, что в точке x_0 существует предел функции f по множеству X_1, причем, согласно условию (5.59), он равен $f(x_0)$:

$$\lim_{\substack{x \to x_0 \\ x \in X_1}} f(x) = f(x_0).$$

С другой стороны, предел функции f при $x \to x_0$ по одноточечному множеству $X_2 = \{x_0\}$, очевидно, равен $f(x_0)$ (предел постоянной равен самой этой постоянной):

$$\lim_{\substack{x \to x_0 \\ x \in X_2}} f(x) = f(x_0).$$

Поэтому, в силу леммы 5 п. 5.8, при $x \to x_0$ у функции f существует и предел, равный $f(x_0)$, по множеству $X = X_1 \cup X_2$:

$$\lim_{x \to x_0} f(x) = f(x_0).$$

Это и означает непрерывность функции f в точке x_0. □

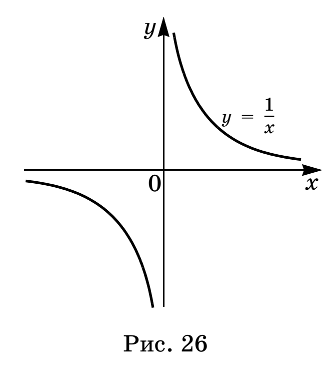


Рис. 26

Точка разрыва функции, не являющейся ее точкой разрыва первого рода, называется *точкой разрыва второго рода*.

Очевидно, что в точках разрыва второго рода по крайней мере один из пределов $\lim\limits_{x \to x_0 - 0} f(x)$ или $\lim\limits_{x \to x_0 + 0} f(x)$ не существует. Здесь под пределом, как обычно, понимается лишь конечный предел.

УПРАЖНЕНИЕ 11. Сформулировать в «позитивном» смысле определение точки разрыва второго рода.

Функции $\operatorname{sign} x$ (см. рис. 19) и $|\operatorname{sign} x|$ (см. рис. 22) имеют в точке $x_0 = 0$ разрыв первого рода, причем у $|\operatorname{sign} x|$ это устранимый разрыв, а функции $\frac{1}{x}$ (рис. 26) и $\sin \frac{1}{x}$ (см. рис. 21) в точке $x_0 = 0$ имеют разрыв второго рода.

5.14. Пределы монотонных функций

Определение 16. *Функция $f : X \to \boldsymbol{R}$, $X \subset \boldsymbol{R}$, называется возрастающей (убывающей) на множестве X, если для любых таких точек $x_1 \in X$ и $x_2 \in X$, что $x_1 < x_2$, выполняется неравенство $f(x_1) \leqslant f(x_2)$ (соответственно неравенство $f(x_1) \geqslant f(x_2)$).*

Возрастающие (убывающие) функции называются иногда *неубывающими* (*невозрастающими*).

Если функция является возрастающей (убывающей) на множестве X, то говорят также, что она возрастает (убывает) на этом множестве.

Если функция f возрастает (убывает) на множестве X, то функция $-f$, получающаяся из f изменением знака у всех ее значений, т. е. $(-f)(x) \stackrel{\text{def}}{=} -f(x)$, $x \in X$, является убывающей (возрастающей) на X функцией.

Возрастающие и убывающие на множестве X функции называются *монотонными* на этом множестве.

ТЕОРЕМА 4. *Пусть функция* $f: X \to \boldsymbol{R}$ *возрастает на множестве* X, $\alpha = \inf X$, $\beta = \sup X$, *причем* $\alpha \notin X$, $\beta \notin X$; *тогда у функции* f *в точке* α *существует предел справа и* $\lim\limits_{x \to \alpha} f(x) = \inf\limits_{x \in X} f(X)$, *а в точке* β *— предел слева и* $\lim\limits_{x \to \beta} f(x) = \sup\limits_{x \in X} f(X)$.

Таким образом, если в условиях теоремы функция f ограничена сверху, то в точке β у нее существует конечный предел слева, а если f не ограничена сверху, то $\lim\limits_{x \to \beta} f(x) = +\infty$.

Аналогично, если функция f ограничена снизу, то в точке α у нее существует конечный предел справа, а если f не ограничена снизу, то $\lim\limits_{x \to \alpha} f(x) = -\infty$.

Подобные утверждения справедливы и для убывающих функций; их можно получить, перейдя от функции f к функции $-f$.

СЛЕДСТВИЕ. *Если функция* f *монотонна на множестве* X, $x_0 \in \boldsymbol{R}$, *множества* $X_<(x_0) \overset{\text{def}}{=} \{x : x \in X,\ x < x_0\}$ *и* $X_>(x_0) \overset{\text{def}}{=} \{x : x \in X,\ x > x_0\}$ *не пусты, а* x_0 *является точкой прикосновения каждого из них, то в точке* x_0 *существуют конечные односторонние пределы*

$$f(x_0 - 0) = \sup_{X_<(x_0)} f(x) \quad \text{и} \quad f(x_0 + 0) = \inf_{X_>(x_0)} f(x), \qquad (5.60)$$

причем в случае возрастающей функции

$$f(x_0 - 0) \leqslant f(x_0 + 0), \qquad (5.61)$$

а в случае убывающей функции

$$f(x_0 - 0) \geqslant f(x_0 + 0). \qquad (5.62)$$

Доказательство теоремы. Пусть $b = \sup\limits_X f(x) \leqslant +\infty$ и $\beta = \sup X$, $\beta \notin X$. Зададим произвольно окрестность $U(b)$ точки b, и пусть η — ее левый конец. Очевидно, $\eta < b$ и поэтому, в силу определения верхней грани функции, существует такая точка $\xi \in X$, что

$$f(\xi) > \eta, \qquad (5.63)$$

причем, в силу условий $\xi \in X$, $\beta = \sup X$ и $\beta \notin X$, имеем $\xi < \beta$.

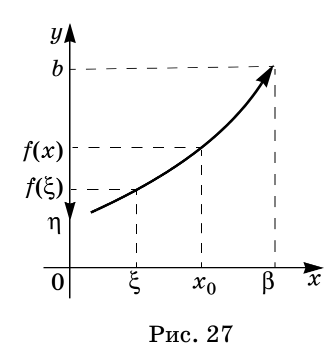


Рис. 27

Обозначим через $U(\beta)$ окрестность точки β, для которой ξ является левым концом (т. е. если β — действительное число, то левым концом интервала $U(\beta, \varepsilon) = (\beta - \varepsilon, \beta + \varepsilon)$, $\varepsilon = \beta - \xi$, а если $\beta = +\infty$, то левым концом бесконечного полуинтервала $(\xi, +\infty]$). Тогда для любой точки

$$x \in X \cap U(\beta) \tag{5.64}$$

имеет место (рис. 27) неравенство $\xi < x$, а следовательно, в силу возрастания функции f, и неравенство $f(\xi) \leqslant f(x)$. Поэтому для всех x, удовлетворяющих условию (5.64), имеем

$$\eta \underset{(5.63)}{<} f(x) \leqslant \sup_X f(x) = b. \tag{5.65}$$

Вспоминая, что точка η является левым концом окрестности $U(b)$ точки b, из (5.65) получим включение $f(x) \in U(b)$. Таким образом, для любой окрестности $U(b)$ точки b существует такая окрестность $U(\beta)$ точки β, что как только $x \in X \cap U(\beta)$, то выполняется включение $f(x) \in U(b)$. Это и означает, что $\lim\limits_{n \to \beta} f(x) = b = \sup\limits_X f(x)$.

Аналогично доказывается, что $\lim\limits_{n \to \alpha} f(x) = \inf\limits_X f(x)$. □

Доказательство следствия. Пусть для определенности функция возрастает на множестве X и x_0 является точкой прикосновения непустых множеств $X_<(x_0)$ и $X_>(x_0)$. Тогда каковы бы ни были точки $x' \in X_<(x_0)$ и $x'' \in X_>(x_0)$, справедливо неравенство $f(x') \leqslant f(x'')$. Поэтому функция f ограничена сверху на множестве $X_<(x_0)$ числом $f(x'')$ и ограничена снизу на множестве $X_>(x_0)$ числом $f(x')$. Следовательно,

$$\sup_{X_<(x_0)} f(x) \leqslant f(x''), \quad \inf_{X_>(x_0)} f(x) \geqslant f(x'). \tag{5.66}$$

В частности, указанные верхние и нижние грани конечны, причем первое из неравенств (5.66) справедливо для любой точки $x'' \in X_>(x_0)$, поэтому, перейдя в его правой части к нижней грани значений функции на множестве $X_>(x_0)$, получим

$$\sup_{X_<(x_0)} f(x) \leqslant \inf_{X_>(x_0)} f(x). \tag{5.67}$$

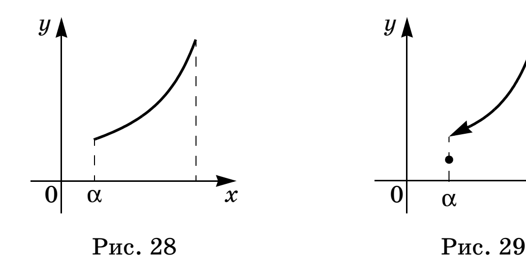


Рис. 28 Рис. 29

Этим завершается доказательство следствия, так как, согласно теореме 4, пределы слева $f(x_0 - 0)$ и справа $f(x_0 + 0)$ существуют, причем

$$f(x_0 - 0) = \sup_{X_<(x_0)} f(x), \quad f(x_0 + 0) = \inf_{X_>(x_0)} f(x),$$

поэтому неравенства (5.61) совпадают с неравенством (5.67). □

Замечание 1. В теореме 4 для возрастающей функции $f : X \to \boldsymbol{R}$ рассмотрены случаи, когда $\inf X = \alpha \notin X$ и $\sup X = \beta \notin X$. Если же, например, $\alpha \in X$, то, как и для произвольной (немонотонной) функции, здесь возможны два случая: предел $\lim\limits_{\substack{x \to \alpha \\ x \in X}} f(x)$ существует, тогда функция f является непрерывной в точке α (рис. 28) или не существует (рис. 29). Аналогичная ситуация имеет место и для точки β.

Замечание 2. Из элементарной математики известно, что функция

$$f(r) = a^r, \ a > 0, \tag{5.68}$$

где r — рациональное число, $r \in \boldsymbol{Q}$, монотонна на множестве всех рациональных чисел $\boldsymbol{Q}$ (см. также п. 2.6*). Для каждого действительного числа x множества рациональных чисел $r < x$, $r > x$ не пусты и x является их точкой прикосновения. Поэтому, согласно следствию теоремы 4, для любого действительного числа x существуют пределы $\lim\limits_{r \to x - 0} a^r$ и $\lim\limits_{r \to x + 0} a^r$, $r \in \boldsymbol{Q}$ (по множеству рациональных чисел $\boldsymbol{Q}$, так как пока у нас показательная функция определена только для рациональных показателей).

В частности, указанные пределы существуют для $x = 0$. Согласно определению предела, их значения равны соответственно значениям пределов последовательностей a^{r_n} при

любых последовательностях аргумента $r_n < 0$ и $r_n > 0$, стремящихся к нулю, $r_n \in \boldsymbol{Q}$, $n = 1, 2, \ldots$. В частности, можно взять $r_n = -\frac{1}{n}$ и $r_n = \frac{1}{n}$, для которых эти пределы были уже вычислены (см. пример 3 в п. 4.9). Тогда получим $\lim\limits_{r \to -0} a^r = \lim\limits_{n \to \infty} a^{-\frac{1}{n}} = 1$, $\lim\limits_{r \to +0} a^r = \lim\limits_{n \to \infty} a^{\frac{1}{n}} = 1$, т. е. односторонние пределы в точке $r = 0$ функции a^r, $r \in \boldsymbol{Q}$, равны, и, следовательно, согласно следствию из теоремы 2 п. 5.9, существует двусторонний предел $\lim\limits_{r \to 0,\, r \neq 0} a^r = 1$. Он совпадает со значением $a^0 = 1$ функции a^r при $r = 0$, а поэтому (см. лемму 3 в п. 5.12) она непрерывна в нуле

$$\lim_{r \to 0} a^r = a^0 = 1, \quad r \in \boldsymbol{Q}, \quad a \in \boldsymbol{R}, \quad a > 0. \tag{5.69}$$

Это равенство будет существенно использовано в п. 7.2 при определении показательной функции a^x для любого действительного числа x, $a > 0$.

Данный пример показывает, что понятие предела по множествам встречается уже в самых простейших ситуациях.

Замечание 3. Из теоремы 4 следует, что всякая монотонная на конечном (бесконечном) интервале функция может иметь только точки разрыва первого рода, причем их множество не более чем счетно (т. е. конечно или счетно).

В самом деле, пусть для определенности функция $f : (a, b) \to \boldsymbol{R}$ возрастает на интервале (a, b), $-\infty \leqslant a < b \leqslant +\infty$. Прежде всего, согласно следствию из теоремы 4, функция f в каждой точке $x_0 \in (a, b)$ имеет конечные пределы слева $f(x_0 - 0)$ и справа $f(x_0 + 0)$, а следовательно, может иметь только разрывы первого рода (определение точки разрыва первого рода см. в п. 5.13), при этом у нее не может быть точек устранимого разрыва. Действительно, если $x_0 \in (a, b)$, то для всех $x' \in (a, x_0)$ и $x'' \in (x_0, b)$, в силу возрастания функции f, справедливо неравенство

$$f(x') \leqslant f(x_0) \leqslant f(x''),$$

откуда

$$\sup_{(a,\, x_0)} f(x) \leqslant f(x_0) \leqslant \inf_{(a,\, x_0)} f(x).$$

Здесь $(a,\ x_0) = \{x \in (a,\ b)\colon x < x_0\}$, а $(x_0,\ b) = \{x \in (a,\ b)\colon x > x_0\}$, поэтому, согласно (5.60), полученное неравенство можно записать в виде

$$f(x_0 - 0) \leqslant f(x_0) \leqslant f(x_0 + 0). \qquad (5.70)$$

Если x_0 — точка устранимого разрыва, т. е. имеет место неравенство $f(x_0 - 0) = f(x_0 + 0)$, то, в силу (5.70), выполняется условие $f(x_0 - 0) = f(x_0) = f(x_0 + 0)$, что означает (см. п. 5.9) непрерывность функции f в точке x_0. Итак, если x_0 — точка разрыва функции f, то $f(x_0 - 0) < f(x_0 + 0)$.

Сопоставим каждой точке разрыва x_0 функции f интервал $(f(x_0 - 0),\ f(x_0 + 0))$ и покажем, что эти интервалы не пересекаются. В самом деле, если x_1 и x_2 — две точки разрыва функции f и, например, $x_1 < x_2$, то $f(x_1 + 0) \leqslant f(x_2 - 0)$. Докажем это. В силу возрастания функции f, для любых точек x' и x'' таких, что $x_1 < x' < x'' < x_2$, справедливо неравенство $f(x') \leqslant f(x'')$. Перейдя в этом неравенстве к пределу при $x' \to$ $\to x_1 + 0$, получим $f(x_1 + 0) \leqslant f(x'')$. Устремляя здесь x'' к x_2 слева: $x'' \to x_2 - 0$, будем иметь $f(x_1 + 0) \leqslant f(x_2 - 0)$, т. е. правый конец интервала $(f(x_1 - 0),\ f(x_1 + 0))$ не больше левого конца интервала $(f(x_2 - 0),\ f(x_2 + 0))$. Отсюда, очевидно, и следует, что указанные интервалы не пересекаются.

Итак, точкам разрыва монотонной функции $f : (a,\ b) \to \boldsymbol{R}$ можно поставить во взаимно однозначное соответствие некоторую систему попарно не пересекающихся интервалов. В каждом таком интервале выберем по одному рациональному числу (такие числа всегда существуют, поскольку множество рациональных чисел всюду плотно на числовой оси (см. замечание в п. 4.11*)). В результате получим взаимно однозначное соответствие между интервалами указанной системы и множеством выбранных рациональных чисел, а следовательно, и между точками разрыва функции и некоторым подмножеством множества рациональных чисел. Но всякое подмножество счетного множества (каким является множество рациональных чисел (см. п. 4.11*)) либо конечно, либо счетно, следовательно, конечно либо счетно множество точек разрыва монотонной функции. □

5.15. Критерий Коши существования предела функции

В настоящем пункте по аналогии со случаем последовательностей будет получено необходимое и достаточное условие того, что функция имеет конечный предел в данной точке x_0, причем это условие будет сформулировано только в терминах значений самой функции, следовательно, значение указанного предела в этом условии не участвует.

Как и раньше, под точкой x_0 понимается либо действительное число, либо одна из бесконечностей ∞, $+\infty$ или $-\infty$, а x_0 является точкой прикосновения множества определения рассматриваемой функции.

ТЕОРЕМА 5 (критерий Коши). *Для того чтобы функция $f: X \to \boldsymbol{R}$ имела в точке x_0 конечный предел, необходимо и достаточно, чтобы для любого $\varepsilon > 0$ существовала такая окрестность $U(x_0)$ точки x_0, что для любых $x' \in U(x_0) \cap X$ и $x'' \in U(x_0) \cap X$ выполнялось бы неравенство*

$$|f(x'') - f(x')| < \varepsilon.$$

Доказательство необходимости. Пусть $f: X \to \boldsymbol{R}$ и $\lim\limits_{x \to x_0} f(x) = a \in \boldsymbol{R}$. Это означает, что для любого $\varepsilon > 0$ существует такая окрестность $U(x_0)$ точки x_0, что для каждого $x \in U(x_0) \cap X$ справедливо неравенство

$$|f(x) - a| < \frac{\varepsilon}{2}. \tag{5.71}$$

Пусть $x' \in U(x_0) \cap X$ и $x'' \in U(x_0) \cap X$; тогда, в силу (5.71), будем иметь

$$|f(x'') - f(x')| = |[f(x'') - a] + [a - f(x')]| \leqslant$$

$$\leqslant |f(x'') - a| + |a - f(x')| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \ \square$$

Доказательство достаточности. Пусть функция $f: X \to \boldsymbol{R}$ такова, что для любого $\varepsilon > 0$ существует такая окрестность $U(x_0)$ точки x_0, что для всех

$$x' \in U(x_0) \cap X, \quad x'' \in U(x_0) \cap X, \tag{5.72}$$

выполняется неравенство

$$|f(x'') - f(x')| < \varepsilon. \tag{5.73}$$

Покажем, что отсюда следует существование у функции f конечного предела в точке x_0. Возьмем какую-либо последовательность $x_n \in X$, $n = 1, 2, \dots$,

$$\lim_{n \to \infty} x_n = x_0 \tag{5.74}$$

и произвольно зададим $\varepsilon > 0$. Для этого ε, согласно сделанному предположению, существует окрестность $U(x_0)$ точки x_0, удовлетворяющая условиям (5.72)—(5.73). В силу же условия (5.74), для этой окрестности $U(x_0)$ существует такое $n_0 \in \boldsymbol{N}$, что при всех $n > n_0$, $n \in \boldsymbol{N}$, имеет место $x_n \in U(x_0)$, а так как $x_n \in X$, то $x_n \in U(x_0) \cap X$, $n = n_0 + 1, n_0 + 2, \dots$. Отсюда, принимая во внимание (5.72)—(5.73), получаем, что для всех $n > n_0$ и всех $m > n_0$ выполняется неравенство

$$|f(x_n) - f(x_m)| < \varepsilon,$$

т. е. числовая последовательность $\{f(x_n)\}$ удовлетворяет условиям критерия Коши для числовых последовательностей (см. п. 4.7) и, следовательно, сходится.

Таким образом, для каждой последовательности $x_n \in X$, $n = 1, 2, \dots$, $\lim\limits_{n \to \infty} x_n = x_0$, последовательность $\{f(x_n)\}$ сходится. Отсюда, как известно (см. лемму 4 в п. 5.6), следует существование конечного предела $\lim\limits_{x \to x_0} f(x)$. □

В том случае, когда x_0 является числом, условие Коши можно сформулировать следующим образом:

для любого $\varepsilon > 0$ *существует такое* $\delta > 0$, *что для любых* $x' \in X$ *и* $x'' \in X$, *удовлетворяющих условиям* $|x' - x_0| < \delta$, $|x'' - x_0| < \delta$, *выполняется неравенство*

$$|f(x'') - f(x')| < \varepsilon.$$

При $x_0 = \infty$ условию Коши можно придать следующий вид:

для любого $\varepsilon > 0$ *существует такое* $\delta > 0$, *что для любых* $x' \in X$ *и* $x'' \in X$, *удовлетворяющих условиям* $|x'| > \delta$, $|x''| > \delta$, *выполняется неравенство* $|f(x'') - f(x')| < \varepsilon$.

Для случая односторонних пределов условие Коши можно перефразировать без термина «окрестность» следующим образом:

для любого $\varepsilon > 0$ *существует такое* η ($\eta < x_0$, *когда рассматривается предел слева, и* $\eta > x_0$, *когда предел справа*), *что для любых* $x' \in X$ *и* $x'' \in X$, *удовлетворяющих условию* $\eta < x' \leqslant x_0$, $\eta < x'' \leqslant x_0$ *или соответственно* $x_0 \leqslant x' < \eta$, $x_0 \leqslant x'' < \eta$, *выполняется неравенство* $|f(x'') - f(x')| < \varepsilon$.

Отметим, что все эти критерии существования предела функции, относящиеся к разным случаям и имеющие разную формулировку, благодаря удачно выбранной терминологии (понятию окрестности) получили единое доказательство.

5.16. Предел и непрерывность композиции функций

Рассмотрим вопрос о существовании конечных и бесконечных пределов композиций функций, каждая из которых имеет соответствующий предел.

Если $f : X \to \boldsymbol{R}$, $g : Y \to \boldsymbol{R}$ и выполнено условие $f(X) \subset Y$, то на множестве X определена композиция $g \circ f$ функций f и g или, как говорят, сложная функция $g[f(x)]$. Рассматриваемые ниже пределы $\lim\limits_{x \to x_0} f(x)$ и $\lim\limits_{y \to y_0} g(y)$ могут быть конечными или бесконечными, а x_0 и y_0 — конечными или бесконечно удаленными точками прикосновения (см. п. 5.4) соответственно множеств X и $f(X)$.

ТЕОРЕМА 6. *Пусть* $f : X \to \boldsymbol{R}$, $g : Y \to \boldsymbol{R}$, $f(X) \subset Y$ *и существуют конечные или бесконечные пределы*

$$\lim_{x \to x_0} f(x) = y_0, \tag{5.75}$$

$$\lim_{y \to y_0} g(y); \tag{5.76}$$

тогда при $x \to x_0$ *существует и предел* (*конечный или бесконечный*) *сложной функции* $g[f(x)]$, *причем*

$$\lim_{x \to x_0} g[f(x)] = \lim_{y \to y_0} g(y).$$

СЛЕДСТВИЕ. *Если* $f : X \to \boldsymbol{R}$, $g : Y \to \boldsymbol{R}$, $f(X) \subset Y$ *и функция* f *непрерывна в точке* $x_0 \in X$, *а функция* g *непрерывна в точке* $y_0 = f(x_0)$, *то сложная функция* $g[f(x)]$ *непрерывна в точке* x_0.

Короче (но менее точно), непрерывная функция от непрерывной функции непрерывна.

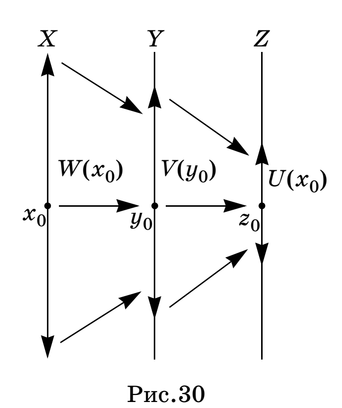


Рис.30

Доказательство теоремы 6. Обозначим значение предела (5.76) через z_0: $\lim\limits_{y \to y_0} g(y) = z_0$ (z_0 — число либо одна из бесконечностей) — и зафиксируем произвольным образом окрестность $U = U(z_0)$ точки z_0. Тогда, согласно определению предела, существует такая окрестность $V = V(y_0)$ точки y, что если

$$y \in Y \cap V(y_0), \tag{5.77}$$

то

$$g(y) \in U(z_0). \tag{5.78}$$

Далее, для полученной окрестности $V(y_0)$, в силу существования предела (5.75), найдется такая окрестность $W = W(x_0)$, что если

$$x \in X \cap W(x_0), \tag{5.79}$$

то

$$f(x) \in V(y_0),$$

а так как $f(x) \in Y$, то

$$f(x) \in Y \cap V(y_0). \tag{5.80}$$

Из выполнения условий (5.79)—(5.80), в силу (5.77)—(5.78), при $y = f(x)$ имеем: если выполнено включение (5.79), то (рис. 30)

$$g\,[f(x)] \in U(z_0).$$

Так как окрестность $U(z_0)$ точки z_0 была произвольна, то это означает, что при $x \to x_0$ у функции $g[f(x)]$ существует предел, равный z_0:

$$\lim_{x \to x_0} g\,[f(x)] = z_0 = \lim_{y \to y_0} g(y). \ \square$$

Утверждение следствия является частным случаем теоремы, когда $y_0 = \lim\limits_{x \to x_0} f(x) = f(x_0)$ и $\lim\limits_{y \to y_0} g(y) = g(y_0)$ (при этих предположениях точки x_0 и y_0 принадлежат соответственно

множествам X и Y, поэтому являются их точками прикосновения):

$$\lim_{x \to x_0} g[f(x)] = \lim_{y \to y_0} g(y) = g(y_0) = g[f(x_0)].$$

З а м е ч а н и е 1. Если $f: X \to \boldsymbol{R}$, $g: X \to \boldsymbol{R}$, существует предел (5.75) и множество Y содержит некоторую окрестность $V(y_0)$ точки y_0:

$$V(y_0) \subset Y, \tag{5.81}$$

то, в силу существования предела (5.75), найдется такая окрестность $W = W(x_0)$ точки x_0, что $f(X \cap W) \subset V(y_0)$, и, следовательно, для сужения f_0 функции f на множестве $X \cap W$ выполнено включение

$$f_0(X \cap W) \subset Y. \tag{5.82}$$

Таким образом, если перейти к сужению f_0 функции f, то при указанном дополнительном предположении (5.81) в условиях теоремы 6 можно не требовать существования композиции функций g и f_0, т. е. выполнения условия $f(X) \subset Y$ — в указанном выше смысле оно выполняется автоматически, а именно имеет место включение (5.82) и поэтому существует композиция $g \circ f_0$.

З а м е ч а н и е 2. Утверждение следствия теоремы 7 можно записать в виде формулы

$$\lim_{x \to x_0} g[f(x)] = g[\lim_{x \to x_0} f(x)], \tag{5.83}$$

из которой видно, что, образно говоря, *операция предельного перехода перестановочна с операцией взятия непрерывной функции*.

В самом деле, левая часть равенства (5.83), в силу непрерывности функции $g[f(x)]$ в точке x_0 (см. следствие теоремы 6), равна $g[(f(x_0)]$. Этому же значению $g[f(x_0)]$ равна и правая часть равенства, но уже в силу непрерывности функции f в той же точке x_0.

З а м е ч а н и е 3. Доказанную в теореме 6 формулу

$$\lim_{x \to x_0} g[f(x)] = \lim_{y \to y_0} g(y), \tag{5.84}$$

где $y_0 = \lim\limits_{x \to x_0} f(x)$, можно рассматривать как *правило замены переменного для вычисления пределов сложных функций*.

Употребляя обозначение композиции функций $g \circ f$, равенство (5.84) можно записать в виде

$$\lim_{x \to x_0} (g \circ f)(x) = \lim_{y \to y_0} g(y).$$

З а м е ч а н и е 4. Пусть $f : X \to \boldsymbol{R}$, $g : Y \to \boldsymbol{R}$, $f(X) \subset Y$ и $\lim\limits_{x \to x_0} f(x) = y_0$. В этом случае, согласно теореме 6, из существования предела (конечного или бесконечного) в правой части равенства (5.84) следует существование соответствующего предела в левой части и равенство этих пределов. Если, кроме того, отображение $f : X \to Y$ является взаимно однозначным отображением множества X н а м н о ж е с т в о Y (т. е. на Y существует однозначная обратная функция f^{-1}) и если $\lim\limits_{y \to y_0} f^{-1}(y) = x_0$, то и наоборот, из существования конечного или бесконечного предела в левой части равенства (5.84) следует существование соответствующего предела в правой части этого равенства.

Таким образом, *при сделанных предложениях предел* (*конечный или бесконечный*) $\lim\limits_{x \to x_0} g[f(x)]$ *существует тогда и только тогда, когда существует* (*конечный или бесконечный*) *предел* $\lim\limits_{y \to y_0} g(y)$, *причем в случае их существования они равны.*

Это утверждение следует из теоремы 6, если ее применить к композиции $(g \circ f) \circ f^{-1}$ функций f^{-1} и $g \circ f$. Согласно теореме 6, из существования пределов $\lim\limits_{y \to y_0} f^{-1}(y) = x_0$ и $\lim\limits_{x \to x_0} (g \circ f)(x)$ следует, что существует предел

$$\lim_{y \to y_0} ((g \circ f) \circ f^{-1})(y) = \lim_{x \to x_0} (g \circ f)(x),$$

но $(g \circ f) \circ f^{-1} = g \circ (f \circ f^{-1}) = g$. Тем самым существует предел $\lim\limits_{y \to y_0} g(y)$, причем $\lim\limits_{y \to y_0} g(y) = \lim\limits_{x \to x_0} (g \circ f)(x)$.

§ 6.

Свойства непрерывных функций на промежутках

6.1. Ограниченность непрерывных функций. Достижимость экстремальных значений

В настоящем параграфе будет изучен ряд важных, находящих многочисленные применения свойств непрерывных функций.

Определение 1. *Функция $f : X \to \boldsymbol{R}$, $X \subset \boldsymbol{R}$, называется непрерывной на множестве X, если она непрерывна по множеству X в каждой его точке* (см. определения 5 и 8 в п. 5.5).

Важным классом непрерывных функций является класс функций, непрерывных на промежутках числовой оси. Начнем его изучение с функций, непрерывных на отрезках. Если функция f непрерывна на отрезке $[a, b]$, то ее непрерывность в точке $x = a$ означает непрерывность справа, а ее непрерывность в точке $x = b$ — непрерывность слева в этих точках.

Наибольшим $\max\limits_X f$ (*наименьшим* $\min\limits_X f$) значением функции $f : X \to \boldsymbol{R}$ называется наибольшее (наименьшее) значение множества всех ее значений (см. п. 3.3). Очевидно, что если у функции f существует наибольшее (наименьшее) значение, то оно является ее верхней (нижней) гранью (см. п. 5.1) $\sup\limits_X f$ (соответственно $\inf\limits_X f$).

ТЕОРЕМА 1 (теорема Вейерштрасса). *Непрерывная на отрезке функция ограничена и принимает на нем наибольшее и наименьшее значение.*

Доказательство. Пусть функция f непрерывна на отрезке $[a, b]$ и пусть $M = \sup\limits_{a \leqslant x \leqslant b} f(x)$; как и всякая верхняя грань непустого множества чисел, M может быть либо конечной, либо бесконечной, равной $+\infty$. Покажем, что $M < +\infty$ и что существует такая точка $x_0 \in [a, b]$, что $f(x_0) = M$.

Выберем какую-либо последовательность таких чисел a_n, $n = 1, 2, \ldots$, что

$$\lim_{n \to \infty} a_n = M, \ a_n < M, \ n = 1, 2, \ldots . \qquad (6.1)$$

Согласно определению верхней грани функции, для каждого a_n, $n = 1, 2, \ldots$, существует такая точка $x_n \in [a, b]$, что

$$f(x_n) > a_n, \ n = 1, 2, \ldots . \qquad (6.2)$$

С другой стороны, поскольку M — верхняя грань функции f, для всех точек $x \in [a, b]$ справедливо неравенство

$$f(x) \leqslant M. \qquad (6.3)$$

Последовательность $\{x_n\}$ ограничена: $a \leqslant x_n \leqslant b$, $n = 1, 2, \ldots$, поэтому по теореме Больцано—Вейерштрасса (см. п. 4.6) из нее можно выделить сходящуюся подпоследовательность $\{x_{n_k}\}$

$$\lim_{k \to \infty} x_{n_k} = x_0. \qquad (6.4)$$

Так как $a \leqslant x_{n_k} \leqslant b$, $k = 1, 2, \ldots$, то (почему?) и $a \leqslant x_0 \leqslant b$, т. е. x_0 — точка отрезка $[a, b]$.

Из неравенств (6.2) и (6.3) следует, что для всех $k = 1, 2, \ldots$ справедливы неравенства

$$a_{n_k} < f(x_{n_k}) \leqslant M. \qquad (6.5)$$

Предел всякой подпоследовательности последовательности, имеющей конечный или бесконечный предел, равен пределу всей последовательности; поэтому из (6.1) имеем $\lim\limits_{k \to \infty} a_{n_k} = M$. Переходя в (6.5) к пределу при $k \to \infty$, получаем

$$\lim_{k \to \infty} f(x_{n_k}) = M. \qquad (6.6)$$

С другой стороны, в силу непрерывности функции f на отрезке $[a, b]$ она непрерывна в точке x_0 этого отрезка и, следовательно, из (6.4) следует, что

$$\lim_{k \to \infty} f(x_{n_k}) = f(x_0). \qquad (6.7)$$

Из формул (6.6) и (6.7) имеем $M = f(x_0)$.

Таким образом, доказано, что верхняя грань M функции f совпадает со значением функции в точке x_0 и, следовательно, конечна. Тем самым функция f ограничена сверху и ее верхняя грань достигается в точке $x_0 \in [a, b]$.

Аналогично доказывается, что непрерывная на отрезке функция ограничена снизу и достигает на нем своей нижней грани. □

Теорема, аналогичная теореме 1, несправедлива для промежутков, не являющихся отрезками; в этом легко убедиться, построив соответствующие примеры. Например, функция $y = 1/x$ непрерывна в каждой точке интервала $(0, 1)$ и вместе с тем не ограничена на нем; функция $y = x$ непрерывна на всей числовой оси и не ограничена на ней.

Отметим еще, что если функция f непрерывна не на отрезке, а на промежутке другого типа и даже, кроме того, ограничена на нем, она, вообще говоря, не имеет наибольшего и наименьшего значения. Например, функция $y = x$ на интервале $(0, 1)$ и $y = \operatorname{arctg} x$ на всей числовой прямой, хотя они непрерывны (непрерывность функции $y = \operatorname{arctg} x$ будет доказана в п. 7.3) и ограничены на указанных промежутках, не достигают своих верхних и нижних граней.

УПРАЖНЕНИЯ. 1. Пусть функция f определена и непрерывна на отрезке $[a, b]$ и $f(x) > 0$ для всех $x \in [a, b]$. Тогда существует такое $c > 0$, что $f(x) > c$ для всех $x \in [a, b]$.

2. Доказать, что функция f, непрерывная на конечном или бесконечном полуинтервале $[a, b)$, $a < b \leqslant +\infty$, и имеющая конечный предел $\lim\limits_{x \to b-0} f(x)$, ограничена на $[a, b)$.

6.2. Промежуточные значения непрерывных функций

ТЕОРЕМА 2 (теорема Больцано—Коши). *Если функция f непрерывна на отрезке $[a, b]$ и $f(a) = A$, $f(b) = B$, то для любого C, заключенного между A и B, существует такая точка, $\xi \in [a, b]$, что $f(\xi) = C$.*

Иначе говоря, *непрерывная на отрезке функция, принимая какие-либо два значения, принимает и любое лежащее между ними значение.*

Доказательство. Пусть для определенности $f(a) = A < B = f(b)$ и $A < C < B$. Разделим отрезок $[a, b]$ точкой x_0 на два равных по длине отрезка; тогда либо $f(x_0) = C$ и, значит, искомая точка $\xi = x_0$ найдена, либо $f(x_0) \neq C$ и тогда на концах одного из полученных отрезков функция f принимает значения, лежащие по разные стороны от числа C, точнее — на левом конце значение, меньшее C, на правом — большее.

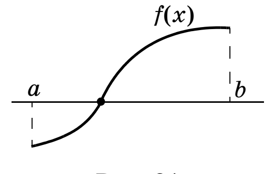


Рис. 31

Обозначим этот отрезок $[a_1, b_1]$ и разделим его снова на два равных по длине отрезка и т. д. В результате либо через конечное число шагов придем к искомой точке ξ, в которой $f(\xi) = C$, либо получим последовательность вложенных отрезков $[a_n, b_n]$, по длине стремящихся к нулю и таких, что

$$f(a_n) < C < f(b_n). \tag{6.8}$$

Пусть ξ — общая точка всех отрезков $[a_n, b_n]$, $n = 1, 2, \ldots$ (см. п. 3.7). Как известно (см. (4.26)), $\xi = \lim\limits_{n \to \infty} a_n = \lim\limits_{n \to \infty} b_n$. Поэтому, в силу непрерывности функции f,

$$f(\xi) = \lim_{n \to \infty} f(a_n) = \lim_{n \to \infty} f(b_n). \tag{6.9}$$

Из (6.8) же получим (см. п. 3.3)

$$\lim_{n \to \infty} f(a_n) \leqslant C \leqslant \lim_{n \to \infty} f(b_n). \tag{6.10}$$

Из (6.9) и (6.10) следует, что $f(\xi) = C$. □

СЛЕДСТВИЕ 1. *Если функция непрерывна на отрезке и на его концах принимает значения разного знака, то на этом отрезке существует хотя бы одна точка, в которой функция обращается в нуль.*

Это следствие — частный случай теоремы (рис. 31).

СЛЕДСТВИЕ 2. *Пусть функция f непрерывна на отрезке $[a, b]$ и $M = \sup f$, $m = \inf f$. Тогда функция f принимает все значения из отрезка $[m, M]$ и только эти значения.*

Для доказательства следствия заметим, что если $M = \sup\limits_{[a, b]} f$, $m = \inf\limits_{[a, b]} f$, то $m \leqslant f(x) \leqslant M$ и, согласно теореме 1, существу-

ют такие точки $\alpha \in [a,\ b]$ и $\beta \in [a,\ b]$, что $f(\alpha) = m$, $f(\beta) = M$. Теперь рассматриваемое следствие непосредственно вытекает из теоремы 2, примененной к отрезку $[\alpha, \beta]$, если $\alpha \leqslant \beta$, или соответственно к отрезку $[\beta, \alpha]$, если $\beta < \alpha$.

Таким образом, *множество всех значений функции, заданной и непрерывной на некотором отрезке, представляет собой также отрезок.*

Отметим, что свойство непрерывных функций принимать все промежуточные значения справедливо для любого промежутка (конечного или бесконечного). Именно: если непрерывная на некотором промежутке функция принимает в двух его точках a и b, причем $a < b$, два каких-то значения, то она принимает и любое промежуточное. В самом деле, согласно теореме 2, рассматриваемая функция заведомо принимает указанное значение в некоторой точке отрезка $[a,\ b]$, который является частью исходного промежутка.

З а м е ч а н и е. Как в теореме 1, так и в теореме 2 было доказано существование точки на данном отрезке, в которой значение рассматриваемой непрерывной функции обладает определенным свойством (в первой теореме в этой точке достигается экстремальное значение, во второй — принимается заданное промежуточное значение). Однако между методами, примененными для доказательства этих утверждений, имеется принципиальное различие. Метод доказательства теоремы 2 дает возможность не только доказать в общем случае существование указанной точки, но и фактически найти ее с любой заданной степенью точности для каждой конкретной функции: нужно разделить отрезок, на котором ищется точка, достаточное число раз пополам, выбирая каждый раз половину согласно правилу, указанному при доказательстве; концы получившегося отрезка и будут приближенными значениями указанной точки.

Метод же доказательства теоремы 1 не позволяет указать способ, с помощью которого для каждой непрерывной на отрезке функции можно было бы найти точки, в которых она принимает экстремальные значения. Это обусловлено тем, что доказательство этой теоремы основано на теореме Больцано—Вейерштрасса, утверждающей лишь

в о з м о ж н о с т ь выделения из каждой ограниченной последовательности сходящейся подпоследовательности. Конкретного метода, или, как это принято говорить, *алгоритма*, для выделения из любой ограниченной последовательности сходящейся подпоследовательности, не существует.

Заметим еще, что при использовании какого-либо алгоритма на практике важно, как быстро он приводит к цели. С этой точки зрения при приближенном решении уравнения $f(x) = 0$ обычно применяется не метод последовательного деления отрезка пополам, а другие алгоритмы, быстрее приводящие к цели (см. § 62 во втором томе).

Задача 7. Доказать, что периодическая непрерывная на всей числовой оси функция, отличная от постоянной, имеет наименьший период. Привести пример периодической функции, определенной на всей числовой оси и отличной от постоянной, которая не имеет наименьшего периода.

6.3. Обратные функции

О п р е д е л е н и е 2. *Функция f, определенная на числовом множестве X, называется строго возрастающей (строго убывающей), если для любых двух чисел $x_1 \in X$ и $x_2 \in X$ таких, что $x_1 < x_2$, выполняется неравенство $f(x_1) < f(x_2)$ (соответственно $f(x_1) > f(x_2)$).*

Функция, строго возрастающая или строго убывающая, называется строго монотонной.

Если функция является строго возрастающей (строго убывающей) на множестве X, то будем также говорить, что она строго возрастает (строго убывает) на этом множестве.

Очевидно, что строго монотонная (возрастающая, убывающая) функция является и просто монотонной (соответственно возрастающей, убывающей) функцией в смысле определения 16 из п. 5.14.

Л Е М М А 1. *Пусть функция f строго возрастает (убывает) на некотором множестве $X \subset \boldsymbol{R}$ и пусть Y — множество ее значений. Тогда обратная функция f^{-1}* (см. п. 1.2*) *является однозначной строго возрастающей (строго убывающей) функцией на множестве Y.*

Доказательство. Пусть для определенности функция f строго возрастает на множестве X. Докажем, что обратная функция однозначна.

Допустим противное. Пусть существует такая точка $y \in Y$, что множество $f^{-1}(y)$ содержит по крайней мере две различных точки x_1 и x_2:

$$x_1 \in f^{-1}(y) \text{ и } x_2 \in f^{-1}(y), \; x_1 \neq x_2,$$

и, следовательно,

$$f(x_1) = f(x_2). \tag{6.11}$$

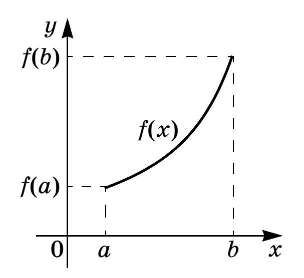
Для двух чисел x_1 и x_2, $x_1 \neq x_2$ справедливо одно из двух неравенств: $x_1 < x_2$ или $x_1 > x_2$; в первом случае, в силу строгого возрастания функции f, имеем $f(x_1) < f(x_2)$, а во втором $f(x_1) > f(x_2)$, т. е. в обоих случаях равенство (6.11) не выполняется. Таким образом, для каждого $y \in Y$ множество $f^{-1}(y)$ состоит в точности из одной точки, т. е. функция f^{-1} однозначна.

Докажем теперь, что функция f^{-1} строго возрастает на множестве Y. Пусть

$$y_1 < y_2, \quad y_1 \in Y, \quad y_2 \in Y \tag{6.12}$$

и пусть $x_1 = f^{-1}(y_1)$, $x_2 = f^{-1}(y_2)$. Следовательно, $y_1 = f(x_1)$, $y_2 = f(x_2)$. Для любых двух чисел x_1 и x_2 справедливо одно из трех соотношений: либо $x_1 > x_2$, либо $x_1 = x_2$, либо $x_1 < x_2$. Если $x_1 > x_2$ или $x_1 = x_2$, то соответственно было бы $y_1 > y_2$ (в силу строгого возрастания функции f) или $y_1 = y_2$ (в силу однозначности), что противоречило бы неравенству (6.12). Таким образом, из неравенства (6.12) следует, что $x_1 < x_2$, а это и означает строгое возрастание функции f^{-1} на множестве Y.

В случае строго убывающей на множестве функции f доказательство можно либо провести аналогичным образом, либо свести к уже рассмотренному случаю рассмотрением функции $-f$, ибо когда функция f строго убывает на множестве X, функция $-f$ строго возрастает на этом множестве. □



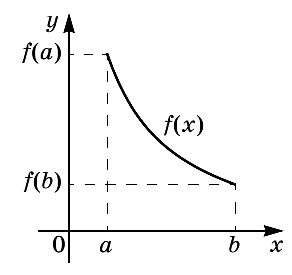


Рис. 32

ТЕОРЕМА 3. *Пусть функция f определена, строго возрастает (строго убывает) и непрерывна на отрезке $[a, b]$; тогда обратная функция f^{-1} определена, однозначна, строго возрастает (строго убывает) и непрерывна на отрезке с концами в точках $f(a)$ и $f(b)$* (рис. 32).

Доказательство. Проведем доказательство теоремы для строго возрастающих функций. Пусть $c = f(a)$, $d = f(b)$.

Покажем, что областью определения обратной функции f^{-1} является сегмент $[c, d]$, или, что то же, $[c, d]$ — множество значений функции f. В самом деле, из возрастания функции f следует, что $f(a) \leqslant f(x) \leqslant f(b)$, т. е. что $f(x) \in [c, d]$ для любого $x \in [a, b]$. С другой стороны, каково бы ни было $y \in [c, d]$, т. е. $f(a) \leqslant y \leqslant f(b)$, согласно теореме 2, существует такая точка $x \in [a, b]$, что $f(x) = y$. Таким образом, все значения заданной функции f лежат на отрезке $[c, d]$, и каждая точка этого отрезка является значением функции f в некоторой точке. Это и означает, что отрезок $[c, d]$ является множеством значений функции f. Отметим, что это утверждение следует также и из следствия 2 теоремы 2, если заметить, что в данном случае

$$c = \min_{[a,\, b]} f(x), \quad d = \max_{[a,\, b]} f(x).$$

В силу леммы, функция f^{-1} однозначная и строго возрастает на отрезке $[c, d]$.

Покажем, наконец, что функция f^{-1} непрерывна на $[c, d]$. Пусть $y_0 \in [c, d]$ и $x_0 = f^{-1}(y_0)$. Пусть $c < y_0 < d$, т. е. y_0 — внутренняя точка отрезка $[c, d]$, тогда, в силу строгого возрастания функции f^{-1}, и $a < x_0 < b$. Зафиксируем некото-

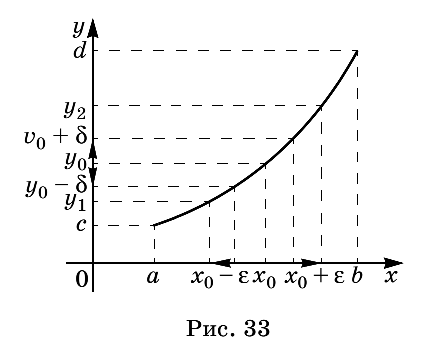


Рис. 33

рое $\varepsilon > 0$. Не ограничивая общности дальнейших рассуждений, можно считать (почему?), что ε таково, что

$$a \leqslant x_0 - \varepsilon < x_0 < < x_0 + \varepsilon \leqslant b. \qquad (6.13)$$

Пусть $y_1 = f(x_0 - \varepsilon)$, $y_2 = f(x_0 + + \varepsilon)$. Тогда из условия (6.13), в силу строгого возрастания f, следует, что $c \leqslant y_1 < y_0 < y_2 \leqslant d$.

Возьмем $\delta > 0$ так, чтобы $y_1 \leqslant y_0 - \delta < y_0 + \delta \leqslant y_2$ (рис. 33). Если теперь выбрать y таким, что $y_0 - \delta < y < y_0 + \delta$, то, тем более, $y_1 < y < y_2$, и, следовательно, в силу строгого возрастания функции f^{-1}, справедливо неравенство

$$x_0 - \varepsilon = f^{-1}(y_1) < f^{-1}(y) < f^{-1}(y_2) = x_0 + \varepsilon.$$

Таким образом, для $\varepsilon > 0$ указано такое $\delta > 0$, что для всех $y \in (y_0 - \delta, y_0 + \delta)$ выполняется неравенство

$$|f^{-1}(y) - f^{-1}(y_0)| < \varepsilon,$$

т. е. функция f^{-1} непрерывна в точке y_0. Если теперь $y_0 = c$ или $y_0 = d$, то аналогичными рассуждениями доказывается, что функция f^{-1} непрерывна справа в точке c и непрерывна слева в точке d. Теорема для строго возрастающих функций доказана.

Напомним, что функция f строго убывает тогда и только тогда, когда функция $-f$ строго возрастает, поэтому справедливость теоремы для строго убывающих функций следует из рассмотренного случая. □

Рассмотрим теперь функции, определенные на интервале.

ТЕОРЕМА 4. *Пусть функция f определена, строго возрастает (строго убывает) и непрерывна на интервале (a, b) (конечном или бесконечном) и пусть*

$$c = \lim_{x \to a + 0} f(x), \; d = \lim_{x \to b - 0} f(x).$$

Тогда обратная функция f^{-1} определена, однозначна, строго возрастает (строго убывает) и непрерывна на интервале (конечном или бесконечном) с концами c и d.

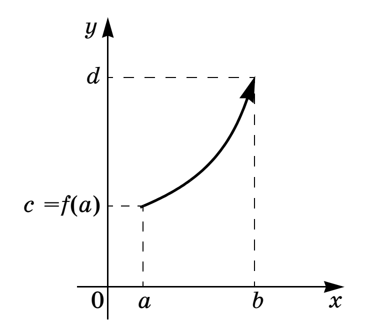
При этом в случае $a = -\infty$ под пределом $\lim\limits_{x \to -\infty + 0} f(x)$ понимается предел $\lim\limits_{x \to -\infty} f(x)$, а в случае $b = +\infty$ под пределом $\lim\limits_{x \to +\infty - 0} f(x)$ — предел $\lim\limits_{x \to +\infty} f(x)$.

Доказательство. Пусть для определенности функция f строго возрастает на интервале (a, b). Покажем, что в этом случае множеством ее значений является интервал (c, d). Действительно, согласно теореме о пределах монотонных функций (см. п. 5.14), имеем $c = \inf\limits_{(a, b)} f$, $d = \sup\limits_{(a, b)} f$ и, следовательно, для любого $x \in (a, b)$ справедливо неравенство $c \leqslant f(x) \leqslant d$. Более того, для всех $x \in (a, b)$ выполняются еще неравенства $f(x) \neq c$, $f(x) \neq d$. В самом деле, если бы, например, существовало такое x_0, что $a < x_0 < b$ и $f(x_0) = c$ (это, очевидно, возможно только тогда, когда нижняя грань c конечна), то при $a < x < x_0$ выполнялось бы неравенство $f(x) < f(x_0) = c$, что противоречило бы тому, что $c = \inf\limits_{(a, b)} f$. Итак, для всех $x \in (a, b)$ выполняются неравенства $c < f(x) < d$. С другой стороны, $c = \inf\limits_{(a, b)} f$, $d = \sup\limits_{(a, b)} f$, поэтому для любого y, $c < y < d$, существуют такие $x_1 \in (a, b)$ и $x_2 \in (a, b)$, что $y_1 = f(x_1)$ и $y_2 = f(x_2)$ удовлетворяют неравенствам

$$c < y_1 < y < y_2 < d.$$

Отсюда следует, что $x_1 < x_2$[1], и поскольку $f(x_1) = y_1$ и $f(x_2) = y_2$, то, в силу теоремы Больцано—Коши о промежуточных значениях непрерывных функций, существует такая точка $x \in [x_1, x_2]$, что $f(x) = y$. Таким образом, для любой точки $y \in (c, d)$ существует такая точка $x \in (a, b)$, что $f(x) = y$.

[1] Случай $x_1 \geqslant x_2$ невозможен, так как тогда бы в силу возрастания функции f выполнялось неравенство $y_1 \geqslant y_2$.



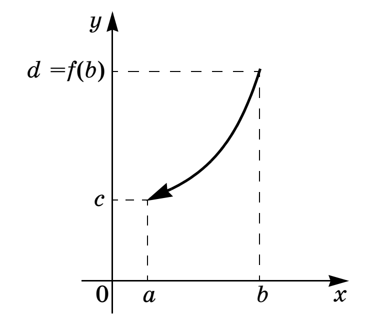


Рис. 34

Тем самым доказано, что действительно множеством значений функции f, или, что то же, множеством определения обратной функции f^{-1}, является интервал $(c,\ d)$. То, что функция f^{-1} однозначна и строго возрастает в интервале $(c,\ d)$, следует из леммы. Ее непрерывность в точке x, $x_1 < x < x_2$, следует из того, что она в силу теоремы 3 непрерывна на отрезке $[x_1,\ x_2] \subset (a,\ b)$.

Как и выше, теорема для строго убывающей функции следует из уже доказанной теоремы о строго возрастающей функции с помощью рассмотрения функции $-f$. □

З а м е ч а н и е. Аналогично доказывается, что если функция строго возрастает и непрерывна на полуинтервале $[a,\ b)$, $-\infty < a < b \leqslant +\infty$, или на $(a,\ b]$, $-\infty \leqslant a < b < +\infty$, то обратная функция определена, строго возрастает и непрерывна на полуинтервале $[c,\ d)$, где $c = f(a)$, $d = \lim\limits_{x \to b - 0} f(x)$, соответственно на $(c,\ d]$, где $c = \lim\limits_{x \to a + 0} f(x)$, $d = f(b)$ (рис. 34).

Случай строго убывающей на полуинтервале функции $f(x)$ можно свести к случаю строго возрастающей функции, рассмотрев функцию $-f(x)$.

П р и м е р. При любом целом положительном n степенна́я функция $y = x^n$ строго возрастает и непрерывна на положительной полуоси $x \geqslant 0$.

Действительно, если $0 \leqslant x_1 < x_2$, то, перемножая n раз эти неравенства, получим $x_1^n < x_2^n$, т. е. что функция $y = x^n$,

$n = 1, 2, \ldots$, строго монотонно возрастает. Для доказательства непрерывности функции $y = x^n$ заметим, что функция $y = f(x) = x$ непрерывна в любой точке $x_0 \in \boldsymbol{R}$. Действительно, в этом случае $y_0 = f(x_0) = x_0$, поэтому $\Delta y = y - y_0 = x - x_0 = \Delta x$. Следовательно, если задано $\varepsilon > 0$, то, взяв $\delta = \varepsilon$, получим, что из условия $|\Delta x| < \delta$ следует $|\Delta y| = |\Delta x| < \delta = \varepsilon$. Это и означает непрерывность функции $y = x$ в точке $x = x_0$. Функция же $y = x^n$ является произведением n одинаковых функций $f(x) = x$ и потому (см. п. 5.10) также непрерывна во всех точках $x \in \boldsymbol{R}$.

Из того, что $\lim\limits_{x \to +\infty} x = +\infty$, очевидно, следует, что $\lim\limits_{x \to +\infty} x^n = +\infty$, $n = 1, 2, \ldots$. Кроме того, в нуле функция $y = x^n$ обращается в нуль. Поэтому, согласно замечанию к теореме 4, множеством значений степенно́й функции $y = x^n$ при $x \geqslant 0$ является неотрицательная полуось $y \geqslant 0$.

Обратной функцией для функции $y^n = x$ является корень n-й степени $\sqrt[n]{y}$, $n = 1, 2, \ldots$. Согласно теореме 4 и доказанным свойствам степенно́й функции $y = x^n$, корень n-й степени $\sqrt[n]{y}$, $n = 1, 2, \ldots$, определен для любого неотрицательного числа y.

Таким образом, из доказанных теорем следует, в частности, *существование и единственность положительного корня n-й степени из любого положительного числа.*

З а м е ч а н и е. Из рассмотренного примера еще раз следует, что любой промежуток содержит иррациональные числа (см. следствие 2 из теоремы 8 в п. 4.11*). Покажем сначала, что число $\sqrt{2}$ (существование которого вытекает из рассмотренного выше примера) является иррациональным.

Допустим противное: пусть существует иррациональное число, равное квадратному корню из двух. Запишем это число в виде несократимой дроби p/q (p и q — взаимно простые натуральные числа): $\sqrt{2} = p/q$. Тогда $p^2 = 2q^2$ и, следовательно, число p делится на 2. Действительно, если бы p было нечетным, т. е. $p = 2k + 1$, $k \in \boldsymbol{N}$, то число $p^2 = (2k + 1)^2 =$

$= 4k^2 + 2k + 1$ также было бы нечетным и равенство $p^2 = 2k^2$ не имело бы места. Итак, $p = 2k$, но тогда $4k^2 = 2q^2$, или $q^2 = 2k^2$. Отсюда, как и выше, следует, что q — четное число. Четность чисел p и q противоречит предположению о несократимости дроби p/q.

Из доказанного, очевидно, следует, что всякое число вида $\frac{m\sqrt{2}}{n}$ (m и n — натуральные) также иррационально. В самом деле, если бы оно было рациональным $\frac{m\sqrt{2}}{n} = \frac{p}{q}$, то и $\sqrt{2}$ оказалось бы рациональным числом: $\sqrt{2} = \frac{np}{mq}$. Отсюда, в свою очередь, следует, что всякий интервал содержит иррациональное число (сравните с п. 4.11*), и притом вида $\frac{m\sqrt{2}}{n}$, m и n — целые.

Действительно, пусть $0 \leqslant a < b$. Выберем натуральное n так, чтобы $\frac{\sqrt{2}}{n} < b - a$, а затем натуральное m так, чтобы

$$\frac{(m-1)\sqrt{2}}{n} \leqslant a < \frac{m\sqrt{2}}{n}.$$

Тогда $a < \frac{m\sqrt{2}}{n} < b$. Если же $a < b \leqslant 0$, то, в силу доказанного, существуют такие целые m и n, что

$$0 \leqslant -b < \frac{m\sqrt{2}}{n} < -a;$$

поэтому $a < -\frac{m\sqrt{2}}{n} < b$.

В случае $a < 0 < b$, согласно доказанному, существуют такие целые m и n, что $a < 0 < \frac{m\sqrt{2}}{n} < b$. □

6.4. Равномерная непрерывность. Модуль непрерывности

Если функция f непрерывна на отрезке, то это означает, что для любой точки x этого отрезка и для любого числа $\varepsilon > 0$ найдется такое число $\delta > 0$ (зависящее от точки x и числа ε), что для всех точек x' отрезка, для которых

$$|x' - x| < \delta, \qquad (6.13)$$

выполняется неравенство

$$|f(x') - f(x)| < \varepsilon. \qquad (6.14)$$

Если число δ можно выбрать не зависящим от точки x, так, чтобы при выполнении условия (6.13) выполнялось условие (6.14), то функция f называется равномерно непрерывной. Сформулируем определение этого важного понятия более подробно.

Определение 3. *Функция f, заданная на отрезке $[a, b]$, называется равномерно непрерывной на нем, если для любого $\varepsilon > 0$ существует такое $\delta > 0$, что для любых двух точек $x \in [a, b]$ и $x' \in [a, b]$ таких, что $|x' - x| < \delta$, выполняется неравенство $|f(x') - f(x)| < \varepsilon$.*

В символической записи определение непрерывности функции на отрезке выглядит следующим образом:

$$\forall\, x\, \forall\, \varepsilon > 0\, \exists\, \delta > 0\, \forall\, x', \quad |x' - x| < \delta\colon |f(x') - f(x)| < \varepsilon,$$

а определение равномерной непрерывности так:

$$\forall\, \varepsilon > 0\, \exists\, \delta > 0\, \forall\, x, x', |x' - x| < \delta\colon |f(x') - f(x)| < \varepsilon. \qquad (6.15)$$

Здесь точки x и x' принадлежат отрезку, на котором рассматривается функция f.

Ясно, что всякая равномерно непрерывная на отрезке функция непрерывна на нем: если в определении равномерной непрерывности зафиксировать точку x, то получится определение непрерывности в этой точке.

Примеры. 1. Функция $f(x) = x$ равномерно непрерывна на всей числовой оси, так как, если задано $\varepsilon > 0$, достаточно взять $\delta = \varepsilon$; тогда если $|x - x'| < \delta$, то, в силу равенств $f(x) = x$, $f(x') = x'$, получим $|f(x) - f(x')| < \varepsilon$.

2. Функция $f(x) = \sin \frac{1}{x}$, $x \neq 0$, будучи непрерывной на своей области определения, т. е. на числовой оси, из которой удалена точка $x = 0$, не будет на ней равномерно непрерывна.

В самом деле, если взять, например, $\varepsilon = 1$, то при любом сколь угодно малом $\delta > 0$ найдутся точки x и x', например точки вида $x = \frac{1}{\pi/2 + 2\pi n}$ и $x' = \frac{1}{3\pi/2 + 2\pi n}$ (n — достаточно большое натуральное число) такие, что $|x - x'| < \delta$, а вместе с тем $|f(x) - f(x')| > 1$.

3. Функция $f(x) = x^2$ не равномерно непрерывна на всей числовой оси $\boldsymbol{R}$. Это следует из того, что для любого фиксированного $h \neq 0$ имеет место равенство

$$\lim_{x \to \infty} [f(x + h) - f(x)] = \lim_{x \to \infty} [(x + h)^2 - x^2] =$$

$$= \lim_{x \to \infty} (2xh + h^2) = \infty.$$

Поэтому если задано $\varepsilon > 0$, то, каково бы ни было $\delta > 0$, зафиксировав $h \neq 0$, $|h| < \delta$, можно так выбрать x, что будет выполняться неравенство $|f(x + h) - f(x)| > \varepsilon$.

ТЕОРЕМА 5 (Кантора). *Функция, непрерывная на отрезке, равномерно непрерывна на нем.*

Доказательство. Докажем теорему от противного. Допустим, что на некотором отрезке $[a, b]$ существует непрерывная, однако не равномерно непрерывная на нем функция f. Это означает (см. (6.15)), что существует такое $\varepsilon_0 > 0$, что для любого $\delta > 0$ найдутся такие точки $x \in [a, b]$ и $x' \in [a, b]$, что $|x' - x| < \delta$, но $|f(x') - f(x) \geqslant \varepsilon_0$. В частности, для $\delta = 1/n$ найдутся такие точки, обозначим их x_n и x_n', что

$$|x_n' - x_n| < \frac{1}{n}, \tag{6.16}$$

но

$$|f(x_n') - f(x_n)| \geqslant \varepsilon_0. \tag{6.17}$$

Из последовательности точек $\{x_n\}$ в силу свойства компактности (см. теорему 4 в п. 4.6) можно выделить сходящуюся подпоследовательность $\{x_{n_k}\}$. Обозначим ее предел x_0:

$$\lim_{k \to \infty} x_{n_k} = x_0. \tag{6.18}$$

Поскольку $a \leqslant x_{n_k} \leqslant b$, $k = 1, 2, \ldots$, то $a \leqslant x_0 \leqslant b$ (см. п. 4.3). Функция f непрерывна в точке x_0, поэтому

$$\lim_{k \to \infty} f(x_{n_k}) \underset{(6.18)}{=} f(x_0). \tag{6.19}$$

Подпоследовательность $\{x_{n_k}'\}$ последовательности $\{x_n'\}$ также сходится к точке x_0, ибо при $k \to \infty$.

$$|x_{n_k}' - x_0| \leqslant |x_{n_k}' - x_{n_k}| + |x_{n_k} - x_0| \underset{(6.16)}{<} \frac{1}{n_k} + |x_{n_k} - x_0| \underset{(6.18)}{\to} 0.$$

Поэтому

$$\lim_{k \to \infty} f(x'_{n_k}) = f(x_0). \qquad (6.20)$$

Из (6.19) и (6.20) следует, что

$$\lim_{k \to \infty} [f(x'_{n_k}) - f(x_{n_k})] = f(x_0) - f(x_0) = 0,$$

а это противоречит условию, что при всех $k = 1, 2, \ldots$ выполняется неравенство

$$|f(x'_{n_k}) - f(x_{n_k})| \underset{(6.17)}{\geqslant} \varepsilon_0 > 0.$$

Полученное противоречие доказывает теорему. □

Условие равномерной непрерывности можно сформулировать в терминах так называемых колебаний функции на отрезках.

Определение 4. *Пусть функция f задана на отрезке $[a, b]$. Тогда величина*

$$\omega(f; [a, b]) = \sup_{x, x' \in [a, b]} |f(x') - f(x)| \qquad (6.21)$$

называется колебанием функции f на отрезке $[a, b]$.

Из двух значений $f(x') - f(x)$ и $f(x) - f(x')$ одно заведомо неотрицательно и, следовательно, не меньше второго, поэтому величина верхней грани в правой части равенства (6.21) не изменится, если вместо абсолютной величины $|f(x') - f(x)|$ разности $f(x') - f(x)$ поставить саму эту разность:

$$\omega(f; [a, b]) = \sup_{x, x' \in [a, b]} [f(x') - f(x)].$$

Справедливо следующее утверждение.

Для того чтобы функция f была равномерно непрерывна на отрезке $[a, b]$, необходимо и достаточно, чтобы для любого $\varepsilon > 0$ существовало такое $\delta > 0$, что каков бы ни был отрезок $[x, x'] \subset [a, b]$ длины меньшей δ: $0 < x' - x < \delta$, выполнялось неравенство

$$\omega(f; [x, x']) < \varepsilon. \qquad (6.22)$$

Действительно, поскольку $x, x' \in [x, x']$, из неравенства (6.22) следует, что $|f(x') - f(x)| < \varepsilon$, поэтому выполняется утверждение (6.15).

Обратно, если справедливо утверждение (6.15), то для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что для любых двух точек x и x' отрезка $[a, b]$, удовлетворяющих условию $|x' - x| < \delta$, имеет место неравенство $|f(x') - f(x)| < \varepsilon/2$.

Пусть для определенности $x < x'$. Для любых двух точек ξ и η отрезка $[x,\ x']$, очевидно, выполняется неравенство $0 < |\eta - \xi| < x' - x < \delta$, следовательно, и неравенство $|f(\eta) - f(\xi)| < \varepsilon/2$. Поэтому для любого отрезка $[x,\ x']$ такого, что $0 < x' - x < \delta$, имеем

$$\omega(f;\, [x,\, x']) = \sup_{\xi,\, \eta \in [x,\, x']} |f(\eta) - f(\xi)| \leqslant \frac{\varepsilon}{2} < \varepsilon. \ \square$$

Часто оказывается удобным еще один подход к понятию равномерной непрерывности, а именно подход, связанный с понятием модуля непрерывности функции.

Определение 5. *Модулем непрерывности* $\omega(\delta;\, f)$ *функции* f, *определенной на отрезке* $[a,\ b]$, *называется функция*

$$\omega(\delta;\, f) = \sup_{|x'' - x'| \leqslant \delta} |f(x'') - f(x')|,\ x',\, x'' \in [a,\ b]. \tag{6.23}$$

Иногда для краткости вместо $\omega(\delta;\, f)$ будем писать просто $\omega(\delta)$. Как и в случае определения колебания функции (6.21), под знаком верхней грани в правой части равенства (6.23) можно не писать знак абсолютной величины разности $|f(x'') - f(x')|$, а брать саму разность — значение верхней грани при этом не изменится:

$$\omega(\delta;\, f) = \sup_{|x'' - x'| \leqslant \delta} [f(x'') - f(x')],\ \ x',\, x'' \in [a,\ b].$$

Очевидно, что $\omega(\delta) \geqslant 0$. Далее, если $0 < \delta_1 < \delta_2$, то

$$\{y : y = f(x'') - f(x'),\, |x'' - x'| \leqslant \delta_1\} \subset$$
$$\subset \{y : y = f(x'') - f(x'),\, |x'' - x'| \leqslant \delta_2\},$$

откуда

$$\sup_{|x'' - x'| \leqslant \delta_1} [f(x'') - f(x')] \leqslant \sup_{|x'' - x'| \leqslant \delta_2} [f(x'') - f(x')],\, x',\, x'' \in [a,\ b],$$

т. е. $\omega(\delta_1) \leqslant \omega(\delta_2)$. Это означает, что модуль непрерывности является возрастающей функцией.

Примеры. 1. Найдем $\omega(\delta)$ для функции $y = x^2$, $-\infty < x < +\infty$.

Для любого $\delta > 0$ и произвольного фиксированного x_0 имеем

$$\omega(\delta;\, x^2) = \sup_{|x'' - x'| \leqslant \delta} ({x''}^2 - {x'}^2) \geqslant$$
$$\geqslant x_0^2 - (x_0 - \delta)^2 = 2x_0\delta - \delta^2. \tag{6.24}$$

Это неравенство верно для всех x_0, и так как при любом фиксированном δ имеем $\lim\limits_{x_0 \to +\infty} (2x_0\delta - \delta^2) = +\infty$, то из (6.24) получаем

$$\omega(\delta; x^2) = +\infty, \ -\infty < x < +\infty.$$

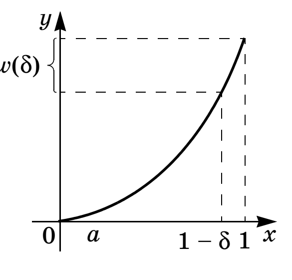


Рис. 35

Найдем теперь модуль непрерывности функции $y = x^2$ на отрезке [0, 1]. Интуитивно ясно, что так как модуль непрерывности $\omega(\delta)$ описывает, согласно определению, наибольший рост функции на отрезке длины δ, то, чтобы получить модуль непрерывности функции, в данном случае следует взять отрезок $[1 - \delta, 1]$, на котором функция $f(x) = x^2$, $0 \leqslant x \leqslant 1$, растет наиболее быстро: модуль непрерывности совпадает (рис. 35) с приращением функции на этом отрезке $\omega(\delta) = f(1) - f(1 - \delta) = 1 - (1 - \delta)^2 = 2\delta - \delta^2$.

Аналитически это проверяется следующим образом. Пусть $0 \leqslant x'' - \delta \leqslant x' \leqslant x'' \leqslant 1$ и, следовательно, $0 \leqslant x'' - x' \leqslant \delta$, тогда, в силу неравенства

$$x''^2 - x'^2 \leqslant x''^2 - (x'' - \delta)^2 = 2x''\delta - \delta^2 \leqslant 2\delta - \delta^2,$$

получим

$$\omega(\delta; x^2) = \sup_{|x'' - x'| \leqslant \delta} (x''^2 - x'^2) \leqslant 2\delta - \delta^2, \tag{6.25}$$

но если взять $x' = 1 - \delta$, $x'' = 1$, то

$$\omega(\delta; x^2) = \sup_{|x'' - x'| \leqslant \delta} (x''^2 - x'^2) \geqslant 1 - (1 - \delta)^2 = 2\delta - \delta^2. \tag{6.26}$$

Из оценок (6.25) и (6.26) следует, что на отрезке [0, 1] имеем $\omega(\delta; x^2) = 2\delta - \delta^2$.

2. Рассмотрим функцию $y = \sin \frac{1}{x}$, $x \neq 0$. С одной стороны,

$$\omega\left(\delta; \sin\frac{1}{x}\right) = \sup_{|x'' - x'| \leqslant \delta} \left|\sin\frac{1}{x''} - \sin\frac{1}{x'}\right| \leqslant$$

$$\leqslant \sup_{|x'' - x'| \leqslant \delta} \left(\left|\sin\frac{1}{x''}\right| + \left|\sin\left|\frac{1}{x'}\right|\right|\right) \leqslant \sup_{|x'' - x'| \leqslant \delta} 2 = 2.$$

С другой стороны, выбрав $x''_n = 1/\left(\frac{\pi}{2} + 2\pi n\right)$, $x'_n = 1/\left(\frac{3}{2}\pi + 2\pi n\right)$ и зафиксировав n так, что $|x'_n| \leqslant \frac{\delta}{2}$, $|x''_n| \leqslant \frac{\delta}{2}$, и поэтому $|x''_n - x'_n| \leqslant |x'_n| + |x''_n| \leqslant \delta$, будем иметь

$$\omega\left(\delta; \sin\frac{1}{x}\right) \geqslant \sin\frac{1}{x''_n} - \sin\frac{1}{x'_n} = 1 + 1 = 2.$$

Из полученных оценок следует, что $\omega\left(\delta; \sin\frac{1}{x}\right) = 2$.

3. Рассмотрим функцию $y = \frac{1}{x}$ на интервале $(0, 1)$.

При любом фиксированном δ, $0 < \delta < 1$, имеем

$$\omega\left(\delta; \frac{1}{x}\right) = \sup_{|x'' - x'| \leqslant \delta}\left(\frac{1}{x'} - \frac{1}{x''}\right) = \sup_{x' \leqslant x'' \leqslant x' + \delta}\left(\frac{1}{x'} - \frac{1}{x''}\right) \geqslant$$

$$\geqslant \frac{1}{x_0} - \frac{1}{x_0 + \delta} = \frac{\delta}{x_0(x_0 + \delta)} \to +\infty \text{ при } x_0 \to +0.$$

Таким образом, $\omega\left(\delta; \frac{1}{\alpha}\right) = +\infty$.

В терминах модуля непрерывности функции ее равномерная непрерывность может быть описана следующим образом.

ТЕОРЕМА 6. *Для того чтобы определенная на отрезке функция была равномерно непрерывна на нем, необходимо и достаточно, чтобы*

$$\lim_{\delta \to 0} \omega(\delta; f) = 0. \qquad (6.27)$$

Доказательство. Пусть функция f равномерно непрерывна на отрезке $[a, b]$, т. е. выполняется условие (6.15), тогда для любого $\varepsilon > 0$ существует такое $\delta_\varepsilon > 0$, что если $x' \in [a, b]$, $x'' \in [a, b]$, $|x'' - x'| < \delta_\varepsilon$, то $|f(x'') - f(x')| < \frac{\varepsilon}{2}$. Отсюда следует, что для любого δ, $0 < \delta < \delta_\varepsilon$, выполняется неравенство

$$\omega(\delta; f) = \sup_{|x'' - x'| \leqslant \delta} |f(x'') - f(x')| \leqslant \frac{\varepsilon}{2} < \varepsilon, \; x', x'' \in [a, b],$$

т. е. если $0 < \delta < \delta_\varepsilon$, то $\omega(\delta; f) < \varepsilon$. Это и означает, что $\lim\limits_{\delta \to 0} \omega(\delta; f) = 0$. Необходимость условия (6.27) доказана.

Докажем его достаточность для равномерной непрерывности. Выполнение условия (6.27) означает, что для любого $\varepsilon > 0$ существует такое $\delta_\varepsilon > 0$, что если $0 < \delta < \delta_\varepsilon$, то

$\omega(\delta; f) < \varepsilon$. Выберем какое-либо из указанных δ. Тогда при $|x'' - x'| < \delta$, $x' \in [a, b]$, $x'' \in [a, b]$ имеем (см. (6.23)):

$$|f(x'') - f(x')| \leqslant \omega(\delta; f) < \varepsilon,$$

т. е. функция f равномерно непрерывна на отрезке $[a, b]$. □

УПРАЖНЕНИЕ. Непрерывная на отрезке $[a, b]$ функция $f(x)$ называется *кусочно-линейной*, если существует такое разбиение отрезка $[a, b]$ на конечное число отрезков $[x_{i-1}, x_i]$,

$$a = x_0 < x_1 < \ldots < x_i < \ldots < x_{n-1} < x_n = b,$$

что функция $f(x)$ линейна на каждом отрезке $[x_{i-1}, x_i]$, $i = 1, 2, \ldots, n$.

Доказать, что всякая непрерывная на отрезке $[a, b]$ функция $F(x)$ может быть с любой степенью точности аппроксимирована кусочно-линейной функцией, т. е. для любого $\varepsilon > 0$ существует такая кусочно-линейная функция $f(x)$, что для всех $x \in [a, b]$ выполняется неравенство $|F(x) - f(x)| < \varepsilon$.

§ 7.

Непрерывность элементарных функций

7.1. Многочлены и рациональные функции

ТЕОРЕМА 1. *Любой многочлен непрерывен в каждой точке числовой оси.*

В самом деле, функция $y = c$, где c — постоянная, непрерывна на всей числовой оси $\boldsymbol{R}$. Это показано в примере 1 п. 5.12.

Функция вида $y = x^n$ также непрерывна при каждом фиксированном $n \in \boldsymbol{N}$ в любой точке $x \in \boldsymbol{R}$. Это показано в п. 6.3 (см. приведенный там пример).

Всякий же многочлен получается из функций вида $y = c$ и $y = x^n$ с помощью их сложения и умножения, а поэтому является непрерывной функцией в каждой точке числовой оси $\boldsymbol{R}$ (см. п. 5.10). □

ТЕОРЕМА 2. *Всякая рациональная функция* $\frac{P(x)}{Q(x)}$ ($P(x)$ *и* $Q(x)$ *— многочлены) непрерывна во всех точках числовой*

оси $\boldsymbol{R}$, *в которых ее знаменатель* $Q(x)$ *не обращается в нуль.*

Это непосредственно следует из того, что многочлены $P(x)$ и $Q(x)$ непрерывны в каждой точке $x \in \boldsymbol{R}$ и частное непрерывных функций также непрерывно во всех точках числовой оси, в которых делитель не обращается в нуль (см. п. 5.10).

Эту теорему удобно использовать при нахождении пределов рациональных функций. Пусть требуется найти $\lim\limits_{x \to x_0} \frac{P(x)}{Q(x)}$. Для этого нужно сначала, если, конечно, это возможно, сократить дробь $\frac{P(x)}{Q(x)}$ на множитель $(x - x_0)^n$ с наибольшим возможным показателем $n \geqslant 1$. Если получившуюся рациональную дробь обозначить $\frac{P_1(x)}{Q_1(x)}$, то (см. п. 5.4)

$$\lim_{x \to x_0} \frac{P(x)}{Q(x)} = \lim_{x \to x_0} \frac{P_1(x)}{Q_1(x)}.$$

Если $Q_1(x_0) \neq 0$, то, в силу теоремы 2, этот предел равен $\frac{P_1(x_0)}{Q_1(x_0)}$, если же $Q_1(x_0) = 0$ (и, значит, $P_1(x_0) \neq 0$, ибо в противном случае дробь $\frac{P_1(x)}{Q_1(x)}$ можно было бы сократить на $x - x_0$), то этот предел равен ∞.

Примеры. 1. $\lim\limits_{x \to 1} \frac{x^2 - 3x + 2}{x^2 - 1} = \lim\limits_{x \to 1} \frac{x - 2}{x + 1} = -\frac{1}{2}$.

2. $\lim\limits_{x \to 1} \frac{x^2 - 3x + 2}{(x - 1)^2} = \lim\limits_{x \to 1} \frac{x - 2}{x - 1} = \infty$.

7.2. Показательная, логарифмическая и степенная функции

Выясним теперь, что понимается под степенью a^x, когда a и x — действительные числа, $a > 0$. Например, что означает выражение $\pi^{\sqrt{2}}$?

Прежде всего напомним свойства степени a^r, где $a > 0$, r — рациональное число: $r = p/q$, p и q — целые, $q \neq 0$ (см. п. 2.6*).

1^0. *Пусть* $r_1 < r_2$. *Если* $a > 1$, *то* $a^{r_1} < a^{r_2}$, *а если* $a < 1$, *то* $a^{r_1} > a^{r_2}$.

2^0. $a^{r_1} \cdot a^{r_2} = a^{r_1 + r_2}$.

3^0. $(a^{r_1})^{r_2} = a^{r_1 r_2}$.

Здесь везде r, r_1 и r_2 — рациональные числа. Вспомним, что $a^0 = 1$. Из свойства 2^0 следует, что $a^r a^{-r} = a^0 = 1$, откуда

$$a^{-r} = \frac{1}{a^r}. \tag{7.1}$$

Далее, из свойства 1^0 и из (7.1) вытекает, что

$$a^r > 0 \tag{7.2}$$

для любого рационального r. Действительно, если $r > 0$ и $a \geqslant 1$, то, в силу 1^0, $a^r \geqslant a^0 = 1 > 0$. Отсюда, согласно (7.1), имеем

$$a^{-r} = \frac{1}{a^r} > 0. \tag{7.3}$$

Аналогично доказывается неравенство $a^r > 0$ при $a < 1$. Отметим еще, что для любых $a > 0$, $b > 0$ и $r \in \boldsymbol{Q}$ имеет место равенство

$$(ab)^r = a^r b^r. \tag{7.4}$$

Определим теперь степень a^x для любого действительного x и $a > 0$.

Определение 1. *Пусть* $a > 0$, x — *произвольное действительное число*, $\boldsymbol{Q}$ — *множество всех рациональных чисел. Положим*

$$a^x = \lim_{r \to x_0,\ r \in \boldsymbol{Q}} a^r. \tag{7.5}$$

Это определение имеет смысл, так как каждая точка числовой оси является точкой прикосновения множества всех рациональных чисел (см. следствие из леммы 1 в п. 4.10). Оно корректно в том смысле, что указанный предел существует, как это доказано ниже, для любого действительного числа $x \in \boldsymbol{R}$. При доказательстве будем использовать определение предела функции по Гейне (см. определение 1 в п. 5.4).

Пусть $a > 0$, $x \in \boldsymbol{R}$, $r_n \in \boldsymbol{Q}$, $n = 1, 2, \ldots$, и $\lim\limits_{n \to \infty} r_n = x$. Покажем, что последовательность $\{a^{r_n}\}$ удовлетворяет условиям критерия Коши (см. п. 4.7) и, значит, является сходящейся. Для этого необходимо оценить разность

$$|a^{r_n} - a^{r_m}| = a^{r_m}|a^{r_n - r_m} - 1|, \quad n \in \boldsymbol{N}, m \in \boldsymbol{N}. \tag{7.6}$$

Последовательность $\{r_n\}$ сходится и, следовательно, ограничена (см. п. 3.4), поэтому существует такое число A, которое без ограничения общности можно считать рациональным (почему?), что $-A < r_n < A$. Отсюда в случае $a \geqslant 1$ имеем $a^{-A} \leqslant \leqslant a^{r_n} < a^A$, а в случае $a < 1$ — соответственно $a^{-A} > a^{r_n} > a^A$, $n = 1, 2, \ldots$, поэтому при любом $a > 0$ существует такое число B, что

$$a^{r_n} \leqslant B, \quad n = 1, 2, \ldots \tag{7.7}$$

($B = a^A$ при $a \geqslant 1$ и $B = a^{-A}$ при $a < 1$), т. е. последовательность $\{a^{r_n}\}$ ограничена сверху числом B.

Далее, согласно замечанию 2 п. 5.14 (см. формулу (5.69)) $\lim\limits_{\substack{r \to 0 \\ n \in 0}} = 1$. Поэтому для любого фиксированного $\varepsilon > 0$ существует такое $\delta = \delta(\varepsilon) > 0$, что для всех рациональных r, удовлетворяющих условию $|r| < \delta$, выполнено неравенство

$$|a^r - 1| < \frac{\varepsilon}{B}. \tag{7.8}$$

Из сходимости же последовательности $\{r_n\}$, в силу критерия Коши (см. п. 3.7), следует, что для найденного $\delta > 0$ существует такой номер n_δ, что для всех $n > n_\delta$ и $m > n_\delta$ выполняется неравенство $|r_n - r_m| < \delta$ и, значит, в силу (7.8), неравенство

$$|a^{r_n - r_m} - 1| < \frac{\varepsilon}{B}. \tag{7.9}$$

Из (7.6), (7.7) и (7.9) вытекает, что для всех $n > n_\delta$ и $m > n_\delta$ справедливо неравенство $|a^{r_n} - a^{r_m}| < \varepsilon$, откуда, в силу критерия Коши, следует, что последовательность $\{a^{r_n}\}$ сходится.

Итак, для любой последовательности рациональных чисел r_n, $n = 1, 2, \dots$, $\lim\limits_{n \to \infty} r_n = x$, последовательность a^{r_n} сходится. Отсюда, согласно лемме 4 из п. 5.6, непосредственно следует существование предела (7.5) функции a^r, $r \in \boldsymbol{Q}$, в точке $x \in \boldsymbol{R}$. Корректность определения a^x доказана.

Определение 1 естественно в том смысле, что в том случае, когда x является рациональным числом r, степень a^x совпадает со значением a^r в ранее известном смысле. В самом деле, если $x = r$ — рациональное число, то в качестве последовательности рациональных чисел r_n, $n = 1, 2, \dots$, сходящейся к числу $x = r$, можно взять $r_n = r$, $n = 1, 2, \dots$. Тогда, согласно определению 1,

$$a^x = \lim_{n \to \infty} a^{r_n} = \lim_{n \to \infty} a^r = a^r. \quad (7.10)$$

Определение 2. *Пусть задано некоторое число $a > 0$. Функция a^x, определенная для всех $x \in \boldsymbol{R}$, называется показательной функцией с основанием a.*

В случае $a = e$ функция e^x обозначается также $\exp x$ и называется *экспонентой*.

Согласно определению 1, $1^x = 1$ для всех действительных x. Поэтому случай $a = 1$ не представляет интереса и в дальнейшем мы не будем его рассматривать.

ТЕОРЕМА 3. *Показательная функция a^x $(a > 0)$ обладает следующими свойствами.*

1^0. *При $a > 1$ она строго возрастает, а при $a < 1$ — строго убывает на всей числовой оси.*

Для любых действительных x и y справедливы следующие равенства:

2^0. $a^x a^y = a^{x+y}$.

3^0. $(a^x)^y = a^{xy}$.

4^0. *Функция a^x непрерывна на всей числовой оси.*

5^0. *Множеством значений функции a^x, $a > 0$, $a \neq 1$, является множество всех положительных чисел.*

Доказательство свойства 1^0. Пусть для определенности $a > 1$ и $x < y$. Существуют (почему?) такие рациональные числа r' и r'', что $x < r' < r'' < y$. Выберем какие-либо последовательности рациональных чисел $\{r'_n\}$ и $\{r''_n\}$ так, чтобы $\lim\limits_{n \to \infty} r'_n = x$, $\lim\limits_{n \to \infty} r''_n = y$ и чтобы $r'_n < r' < r'' < r''_n$ для всех $n = 1, 2, \ldots$. Тогда

$$a^{r'_n} < a^{r'} < a^{r''} < a^{r''_n}; \tag{7.11}$$

перейдя к пределу при $n \to \infty$, получим

$$a^x \leqslant a^{r'} < a^{r''} \leqslant a^y.$$

Таким образом, если $x < y$, то $a^x < a^y$, что и означает строгое возрастание функции a^x при $a > 1$.

Случай $a < 1$ рассматривается аналогичным образом. □

Доказательство свойства 2^0. Пусть $\{r'_n\}$ и $\{r''_n\}$ — такие последовательности рациональных чисел, что $\lim\limits_{n \to \infty} r'_n = x$, $r''_n = y$ и, значит, $\lim\limits_{n \to \infty} (r'_n + r''_n) = x + y$ (см. п. 3.9). Тогда, в силу определения показательной функции,

$$a^{x+y} = \lim_{n \to \infty} a^{r'_n + r''_n} = \lim_{n \to \infty} a^{r'_n} a^{r''_n} =$$

$$= \lim_{n \to \infty} a^{r'_n} \lim_{n \to \infty} a^{r''_n} = a^x a^y. \ \square$$

Прежде чем переходить к доказательству следующих свойств, заметим, что из свойства 2^0 следует, что для любого действительного x справедливо равенство $a^x a^{-x} = a^0 = 1$, поэтому $a^{-x} = 1/a^x$.

Доказательство свойства 4^0. (Свойство 3^0 будет доказано после доказательства свойства 4^0.) Прежде всего отметим, что

$$\lim_{\substack{x \to 0 \\ x \in \boldsymbol{R}}} a^x = 1. \tag{7.12}$$

Действительно, в силу доказанной монотонности фукнции a^x, $a > 0$ (свойство 1^0), существуют односторонние пределы $\lim\limits_{x \to -0} a^x$ и $\lim\limits_{x \to +0} a^x$. Рассуждая аналогично доказательству формулы (5.69) (только не предполагая, что x — рациональное число, а рассматривая любые действительные числа), получаем

$$\lim_{x \to -0} a^x = \lim_{r \to -0,\, r \in Q} a^r = 1 = \lim_{r \to +0,\, r \in Q} a^r = \lim_{x \to +0} a^x,$$

отсюда и следует равенство (7.12).

Пусть теперь $x \in \boldsymbol{R}$, x фиксировано, $y = a^x$ и

$$\Delta y = a^{x + \Delta x} - a^x = a^x a^{\Delta x} - a^x = a^x(a^{\Delta x} - 1).$$

Тогда, в силу (7.12), $\lim\limits_{\Delta x \to 0} a^{\Delta x} = 1$ и поэтому

$$\lim_{\Delta x \to 0} \Delta y = a^x \lim_{\Delta x \to 0} (a^{\Delta x} - 1) = 0,$$

а это и означает непрерывность функции a^x в точке x. □

Доказательство свойства 3^0. Пусть сначала $y = p$ — целое положительное число; тогда применив p раз свойство 2^0, получим

$$(a^x)^p = \underbrace{a^x \cdot a^x \ldots a^x}_{p \text{ раз}} = a^{\overbrace{x + x + \ldots + x}^{p \text{ раз}}} = a^{xp}. \qquad (7.13)$$

Пусть, далее, $y = 1/q$, где q — целое положительное число. Покажем, что $(a^x)^{1/q} = a^{x/q}$, т. е. что $a^{x/q}$ является корнем q-й степени из числа a^x. Для этого, согласно определению корня, надо доказать, что $(a^{x/q})^q = a^x$; это следует из равенства (7.13).

Пусть теперь $y = p/q$, p и q — натуральные, тогда, согласно уже доказанному,

$$(a^x)^{p/q} = [(a^x)^p]^{1/q} = (a^{xp})^{1/q} = a^{xp/q}.$$

Если же $y = -p/q$, то

$$(a^x)^{-p/q} = \frac{1}{(a^x)^{p/q}} = \frac{1}{a^{x\,p/q}} = a^{-xp/q}.$$

Наконец, очевидно, что $(a^x)^0 = 1 = a^0$. Таким образом доказано, что для любого действительного x и любого рационального r

$$(a^x)^r = a^{xr}. \tag{7.14}$$

Пусть теперь задано еще одно действительное число y. Рассмотрим произвольную последовательность $\{r_n\}$ рациональных чисел, сходящуюся к y. Тогда, в силу (7.14), для всех $n = 1, 2, \ldots$ будем иметь

$$(a^x)^{r_n} = a^{x\,r_n}. \tag{7.15}$$

Поскольку $\lim\limits_{n \to \infty} xr_n = xy$, то, согласно доказанной выше непрерывности функции a^x, имеем

$$\lim_{n \to \infty} a^{x\,r_n} = a^{xy}. \tag{7.16}$$

С другой стороны, в силу определения показательной функции,

$$\lim_{n \to \infty} (a^x)^{r_n} = (a^x)^y. \tag{7.17}$$

Переходя к пределу в равенстве (7.15) при $n \to \infty$, из (7.16) и (7.17) получим рассматриваемое свойство для любых $x, y \in \boldsymbol{R}$. □

Заметим, что из свойств 2^0 и 3^0 следует, что

$$\left(\frac{1}{a}\right)^x = \frac{1}{a^x}, \; a > 0, \; x \in \boldsymbol{R}.$$

Действительно,

$$\left(\frac{1}{a}\right)^x = (a^{-1})^x = a^{-x} = \frac{1}{a^x}.$$

Доказательство свойства 5^0. Пусть снова для определенности $a > 1$. Для того чтобы доказать, что множеством значений функции a^x является множество всех положительных чисел, т. е. бесконечный интервал $(0, +\infty)$, в силу ее непрерывности и строгого возрастания на всей числовой оси, согласно теореме 4 п. 6.3, достаточно показать, что

$$\lim_{x \to +\infty} a^x = +\infty, \quad \lim_{x \to -\infty} a^x = 0. \tag{7.18}$$

В силу монотонности функции a^x, пределы (конечные или бесконечные) $\lim\limits_{x \to +\infty} a^x$ и $\lim\limits_{x \to -\infty} a^x$ существуют, следовательно, достаточно доказать, что

$$\lim_{n \to \infty} a^{x_n} = +\infty, \quad \lim_{n \to \infty} a^{x'_n} = 0$$

для каких-либо произвольных фиксированных последовательностей $x_n \to +\infty$ и $x'_n = -\infty$, например для последовательностей $x_n = n$, $x''_n = -n$, $n = 1, 2, \ldots$.

По предположению, $a > 1$, т. е. $a = 1 + \alpha$, где $\alpha > 0$. Поэтому, согласно неравенству Бернулли (см. лемму в п. 4.9), $a^n = (1 + \alpha)^n > n\alpha$, и так как $\lim\limits_{n \to \infty} n\alpha = +\infty$, то и $\lim\limits_{n \to \infty} a^n = +\infty$. Отсюда

$$\lim_{n \to \infty} a^{-n} = \frac{1}{\lim\limits_{n \to \infty} a^n} = 0.$$

Тем самым равенство (7.18) при $a > 1$ доказано.

Если теперь $0 < a < 1$, то $b = 1/a > 1$ и

$$\lim_{x \to +\infty} a^x = \lim_{x \to +\infty} \left(\frac{1}{b}\right)^x = \frac{1}{\lim\limits_{x \to +\infty} b^x} = 0,$$

$$\lim_{x \to -\infty} a^x = \frac{1}{\lim\limits_{x \to -\infty} b^x} = +\infty. \ \square$$

З а м е ч а н и е 1. Множество всех значений функции a^x, $a > 0$, $a \neq 0$, составляет множество всех положительных действительных чисел, поэтому, в частности, при любом $x \in \boldsymbol{R}$ имеет место неравенство $a^x > 0$.

З а м е ч а н и е 2. Если $a > 0$, $b > 0$, то для любого $x \in \boldsymbol{R}$ справедливо равенство

$$(ab)^x = a^x b^x.$$

Действительно, если $r_n \to x$, $r_n \in \boldsymbol{Q}$, $n = 1, 2, \ldots$, то

$$(ab)^x = \lim_{n \to \infty} (ab)^{r_n} = \lim_{n \to \infty} a^{r_n} b^{r_n} =$$

$$= \lim_{n \to \infty} a^{r_n} \lim_{n \to \infty} b^{r_n} = a^x b^x. \ \square$$

УПРАЖНЕНИЕ. Пусть $a > 0$, $b > 0$. Доказать, что для любого $x \in \boldsymbol{R}$ имеет место равенство $\left(\frac{a}{b}\right)^x = \frac{a^x}{b^x}$.

З а м е ч а н и е 3. Если r — рациональное число и $r > 0$, то $0^r = 0$, и, следовательно, для любого действительного числа $x > 0$ существует предел $\lim\limits_{r \to x,\, r \in \boldsymbol{Q}} 0^r = 0$. Поэтому при $x > 0$ определение (7.5) можно распространить и на случай $a = 0$, причем будет иметь место равенство $0^x = 0$, $x > 0$.

Отметим, что в области действительных чисел возведению нуля в неположительную степень: 0^x, $x \leqslant 0$ — нельзя приписать смысла.

Пусть a — положительное число, не равное единице. Из элементарной математики известно, что операция, обратная возведению в степень и ставящая в соответствие данному числу $x > 0$ такое число y, что $a^y = x$ (если, конечно, указанное y существует), называется логарифмированием по основанию a. Число y называется *логарифмом* числа x по основанию a и обозначается через $\log_a x$. Таким образом, по определению,

$$a^{\log_a x} = x \ (a > 0, a \neq 1). \tag{7.19}$$

Определение 3. *Функция, ставящая в соответствие каждому числу x его логарифм $\log_a x$ по основанию a* ($a > 0$, $a \neq 1$), *если этот логарифм существует, называется логарифмической функцией $y = \log_a x$.*

Логарифмическая функция по основанию 10 обозначается символом lg, а по основанию e — символом ln; $\ln x$ называется *натуральным* логарифмом числа x.

ТЕОРЕМА 4. *Функция $y = \log_a x$, $a > 0$, $a \neq 1$, определена на множестве всех $x > 0$ и является на этом множестве строго монотонной (возрастающей при $a > 1$ и убывающей при $a < 1$) непрерывной функцией. Она имеет следующие свойства:*

1^0. $\log_a x_1 x_2 = \log_a x_1 + \log_a x_2$, $x_1 > 0, x_2 > 0$.

2^0. $\log_a x^\alpha = \alpha \log_a x$, $x > 0$, $\alpha \in \boldsymbol{R}$.

Доказательство. Множеством значений функции a^x, $a > 0$, $a \neq 1$, является множество всех положительных чисел $(0, +\infty)$, поэтому это же множество является и множеством определения обратной функции, которой в данном случае является функция $\log_a x$ (см. (7.19)). Этим, в частности, доказано существование логарифма любого положительного числа. Остальные утверждения теоремы 4 непосредственно следуют из теоремы 4 п. 6.3 и теоремы 3 настоящего параграфа.

Например, покажем, как свойство 1^0 вытекает из свойств показательной функции, указанных в теореме 3. Положим $y_1 = \log_a x_1$, $y_2 = \log_a x_2$. Согласно определению логарифма, это означает, что $x_1 = a^{y_1}$, $x_2 = a^{y_2}$. Отсюда (см. свойство 2^0 показательной функции в теореме 3) имеем

$$x_1 x_2 = a^{y_1} a^{y_2} = a^{y_1 + y_2},$$

и, следовательно, снова, согласно определению логарифма,

$$\log_a x_1 x_2 = y_1 + y_2 = \log_a x_1 + \log_a x_2. \ \square$$

Определение 4. *Пусть задано действительно число* α. *Функция* x^α, *определенная для всех* $x > 0$, *называется степенно́й функцией с показателем* α.

ТЕОРЕМА 5. *Степенная функция* x^α, $x > 0$, *непрерывна, строго возрастает при* $\alpha > 0$ *и строго убывает при* $\alpha < 0$.

Действительно, из определения логарифма имеем $x = e^{\ln x}$, а поэтому $x^\alpha = e^{\alpha \ln x}$, т. е. x^α есть композиция показательной функции e^u и логарифмической функции, умноженной на постоянную: $u = \alpha \ln x$. Показательная и логарифмическая функции непрерывны (см. теоремы 3 и 4), поэтому, в силу теоремы о непрерывности композиции непрерывных функций (см. п. 5.2), функция x^α также непрерывна.

Функция $\ln x$ строго возрастает, поэтому функция $x^\alpha = e^{\alpha \ln x}$ строго возрастает при $\alpha > 0$ и строго убывает при $\alpha < 0$. $\square$

При рассмотрении функции x^α предполагалось, что $x > 0$. В точке $x = 0$ функция x^α определена не для всех значений показателя α. Если $\alpha > 0$, то существует предел $\lim\limits_{x \to +0} x^\alpha = \lim\limits_{x \to +0} e^{\alpha \ln x} = 0$, так как $\lim\limits_{x \to +0} \ln x = -\infty$. По определению полагают (см. замечание 3) $0^\alpha = 0$, $\alpha > 0$. При этом определении (7.20) функция x^α оказывается непрерывной справа в точке $x = 0$ при любом $\alpha > 0$:

$$\lim_{x \to +0} x^\alpha = 0 = 0^\alpha. \tag{7.20}$$

Функция x^α может оказаться определенной при некоторых $\alpha \neq 0$ и для $x < 0$, например, $x^{\pm n}$, $x^{\pm 1/(2n-1)}$, $n \in \boldsymbol{N}$.

Степенная функция x^α непрерывна во всех точках, в которых она определена. Это следует из того, что если она определена при $x < 0$, то является четной или нечетной функцией. Если же четная или нечетная функция непрерывна при $x > 0$, то она непрерывна и при $x < 0$ (почему?). Если же в точке $x = 0$ четная или нечетная функция непрерывна справа и равна нулю, то она просто непрерывна в этой точке (почему?). Этот случай имеет место при $\alpha > 0$ (см. (7.20)).

7.3. Тригонометрические и обратные тригонометрические функции

Перейдем к вопросу о непрерывности тригонометрических функций. При этом не будем приводить строгих аналитических определений этих функций (как это было сделано выше в случае показательной функции), а используем их геометрическое определение, известное из элементарной математики. Всюду в дальнейшем x — действительное число, а под $\sin x$, $\cos x$, $\operatorname{tg} x$, $\operatorname{ctg} x$ будем подразумевать значение соответствующей тригонометрической функции угла, радианная мера которого равна x.

ЛЕММА. *При любом действительном x справедливо неравенство*

$$|\sin x| \leqslant |x|.$$

Доказательство. Рассмотрим окружность радиуса R с центром в точке O. Пусть радиус OB образует угол x, $0 \leqslant x \leqslant \frac{\pi}{2}$, с радиусом OA, а радиус OB_1 симметричен радиусу OB относительно OA (рис. 36).

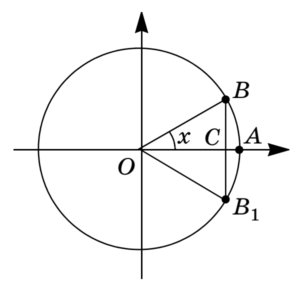


Рис. 36

Опустим из точки B перпендикуляр BC на радиус OA. Тогда $BC = R \sin x$ и, так как $BC = CB_1$, имеем $BB_1 = 2R \sin x$. Как известно, длина дуги BAB_1 равна $2Rx$. Длина отрезка, соединяющего две точки, не превышает длины дуги окружности, соединяющей те же точки, значит, $2R \sin x \leqslant 2Rx$, т. е. $\sin x \leqslant x$.

Если теперь $-\frac{\pi}{2} \leqslant x < 0$, то $0 < -x \leqslant \frac{\pi}{2}$, и поэтому, в силу доказанного, $\sin(-x) \leqslant -x$, но в данном случае $\sin(-x) = |\sin x|$ и $-x = |x|$, следовательно, $|\sin x| \leqslant |x|$. Таким образом, если $|x| \leqslant \frac{\pi}{2}$, то $|\sin x| \leqslant |x|$. Если же $|x| > \frac{\pi}{2}$, то

$$|\sin x| \leqslant 1 < \frac{\pi}{2} < |x|. \ \square$$

ТЕОРЕМА 6. *Функции $y = \sin x$, $y = \cos x$ непрерывны на всей числовой оси.*

СЛЕДСТВИЕ. *Функции $y = \operatorname{tg} x$ и $y = \operatorname{ctg} x$ непрерывны при всех x, при которых $\cos x$ (соответственно $\sin x$) не обращаются в нуль.*

Доказательство. Так как $|\sin \alpha| \leqslant 1$, $|\cos \alpha| \leqslant 1$ при любом α и, в силу леммы, $\left|\sin \frac{\Delta x}{2}\right| \leqslant \frac{1}{2}|\Delta x|$, то

$$|\sin(x + \Delta x) - \sin x| \leqslant 2\left|\sin \frac{\Delta x}{2}\right|\left|\cos\left(x + \frac{\Delta x}{2}\right)\right| \leqslant |\Delta x|,$$

$$|\cos(x + \Delta x) - \cos x| \leqslant 2\left|\sin \frac{\Delta x}{2}\right|\left|\sin\left(x + \frac{\Delta x}{2}\right)\right| \leqslant |\Delta x|.$$

Отсюда следует, что при $\Delta x \to 0$ левые части неравенства также стремятся к нулю. Это и означает непрерывность функций $\sin x$ и $\cos x$.

Непрерывность $\operatorname{tg} x = \frac{\sin x}{\cos x}$ и $\operatorname{ctg} x = \frac{\cos x}{\sin x}$ в точках, в которых знаменатели не обращаются в нуль, следует из непре-

рывности $\sin x$ и $\cos x$ и теоремы о частном непрерывных функций (см. п. 5.10). □

ТЕОРЕМА 7. *Каждая из обратных тригонометрических функций* $\arcsin x$, $\arccos x$, $\operatorname{arctg} x$ *и* $\operatorname{arcctg} x$ *непрерывна в области своего определения.*

Это сразу следует из теорем 3 и 4 в § 6 и из непрерывности и строгой монотонности функций $\sin x$ на отрезке $[-\pi/2, \pi/2]$, $\cos x$ на отрезке $[0, \pi]$, $\operatorname{tg} x$ на интервале $(-\pi/2, \pi/2)$ и $\operatorname{ctg} x$ на интервале $(0, \pi)$.

7.4. Непрерывность элементарных функций

Непрерывность основных элементарных функций, доказанная в этом параграфе, позволяет доказать теорему и о непрерывности произвольных элементарных функций.

ТЕОРЕМА 8. *Всякая элементарная функция непрерывна во всех точках своего множества определения.*

Доказательство. Согласно определению, всякая элементарная функция получается из основных элементарных функций с помощью конечного числа арифметических операций и композиций (см. п. 5.3), поэтому ее непрерывность на множестве определения сразу следует из непрерывности основных элементарных функций на множествах их определения (теоремы 1—7), из свойств пределов функций, связанных с арифметическими действиями над функциями (см. п. 5.10), и непрерывности композиции непрерывных функций (см. п. 5.16). □

§ 8.

Сравнение функций. Вычисление пределов

8.1. Некоторые замечательные пределы

В этом пункте вычисляются пределы, которые неоднократно будут встречаться в дальнейшем.

ЛЕММА 1.

$$\lim_{x \to 0} \frac{\sin x}{x} = 1. \qquad (8.1)$$

Доказательство. Рассмотрим круг радиуса R с центром в точке O. Пусть радиус OB образует угол x, $0 < x < \pi/2$, с радиусом OA. Соединим точки A и B отрезком и восставим из точки A перпендикуляр к радиусу OA до пересечения в точке C с продолжением радиуса OB (рис. 37). Тогда площадь треугольника AOB равна $\frac{1}{2}R^2 \sin x$, площадь сектора AOB равна $\frac{1}{2}R^2 x$, а

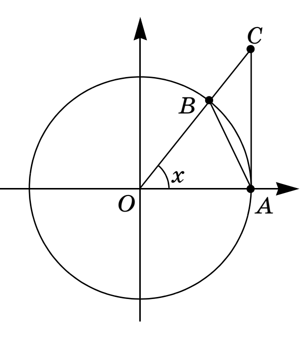


Рис. 37

площадь треугольника AOC равна $\frac{1}{2}R^2 \operatorname{tg} x$. Треугольник AOB является частью сектора AOB, который, в свою очередь, является частью треугольника AOC; поэтому

$$\frac{1}{2}R^2 \sin x < \frac{1}{2}R^2 x < \frac{1}{2}R^2 \operatorname{tg} x,$$

откуда $\sin x < x < \operatorname{tg} x$, следовательно,

$$1 < \frac{x}{\sin x} < \frac{1}{\cos x}$$

или, заменяя величины им обратными,

$$\cos x < \frac{\sin x}{x} < 1. \qquad (8.2)$$

Заметим, что в силу четности функций $\cos x$ и $\frac{\sin x}{x}$ неравенство (8.2) справедливо и при $-\pi/2 < x < 0$.

Так как функция $\cos x$ непрерывна и $\cos 0 = 1$, то из (8.2) при $x \to 0$ следует (см. п. 5.10) равенство (8.1). □

СЛЕДСТВИЕ 1.

$$\lim_{x \to 0} \frac{\operatorname{tg} x}{x} = 1. \qquad (8.3)$$

В самом деле,

$$\lim_{x \to 0} \frac{\operatorname{tg} x}{x} = \lim_{x \to 0} \frac{\sin x}{x} \lim_{x \to 0} \frac{1}{\cos x} = 1.$$

СЛЕДСТВИЕ 2.

$$\lim_{x \to 0} \frac{\arcsin x}{x} = 1. \qquad (8.4)$$

Функция $y = \sin x$ строго монотонна и непрерывна на отрезке $[-\pi/2, \pi/2]$, поэтому обратная функция $x = \arcsin y$ также строго монотонна и непрерывна на отрезке $[-1; 1]$. Так как $\sin 0 = 0$, то

$$\lim_{x \to 0} y = \lim_{x \to 0} \sin x = 0 \quad \text{и} \quad \lim_{y \to 0} x = \lim_{x \to 0} \arcsin x = 0.$$

Чтобы вычислить предел (8.4), применим правило замены переменного для пределов непрерывных функций (см. теорему 6 в п. 5.16). Положив $x = \sin y$, имеем

$$\lim_{x \to 0} \frac{\arcsin x}{x} = \lim_{y \to 0} \frac{\arcsin(\sin y)}{\sin y} = \lim_{y \to 0} \frac{y}{\sin y} = 1.$$

СЛЕДСТВИЕ 3.

$$\lim_{x \to 0} \frac{\operatorname{arctg} x}{x} = 1. \qquad (8.5)$$

Это равенство получается аналогично предыдущему из (8.3).

ЛЕММА 2.

$$\lim_{x \to 0} (1 + x)^{\frac{1}{x}} = e. \qquad (8.6)$$

Ранее (см. п. 4.5) было доказано, что

$$\lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n = e, \qquad (8.7)$$

где $n = 1, 2, \ldots$. Отсюда, в силу леммы п. 4.3, следует, что для любой последовательности $\{n_k\}$ натуральных чисел такой, что

$$\lim_{k \to \infty} n_k = +\infty, \qquad (8.8)$$

имеем

$$\lim_{k \to \infty} \left(1 + \frac{1}{n_k}\right)^{n_k} = e. \qquad (8.9)$$

Пусть теперь последовательность $\{x_k\}$ такая, что $\lim\limits_{k \to \infty} x_k = +0$, т. е.

$$\lim_{k \to \infty} x_k = 0 \text{ и } x_k > 0. \qquad (8.10)$$

Покажем, что $\lim\limits_{k \to \infty} (1 + x_k)^{1/x_k} = e$. При этом без ограничения общности можно считать, что $x_k < 1$, $k = 1, 2, \ldots$ (почему?). Для всякого x_k найдется такое натуральное n_k, что $n_k + 1 > \frac{1}{x_k} \geqslant n_k$ и, следовательно, $\frac{1}{n_k + 1} < x_k \leqslant \frac{1}{n_k}$, причем в силу (8.10), $\lim\limits_{k \to \infty} n_k = +\infty$. Поэтому имеем

$$\left(1 + \frac{1}{n_k + 1}\right)^{n_k} < (1 + x_k)^{1/x_k} < \left(1 + \frac{1}{n_k}\right)^{n_k + 1}. \qquad (8.11)$$

Замечая, что, согласно (8.9),

$$\lim_{k \to \infty} \left(1 + \frac{1}{n_k + 1}\right)^{n_k} =$$

$$= \lim_{k \to \infty} \frac{\left(1 + \frac{1}{n_k + 1}\right)^{n_k + 1}}{1 + \frac{1}{n_k + 1}} = \frac{\lim\limits_{k \to \infty}\left(1 + \frac{1}{n_k + 1}\right)^{n_k + 1}}{\lim\limits_{k \to \infty}\left(1 + \frac{1}{n_k + 1}\right)} = e$$

и

$$\lim_{k \to \infty} \left(1 + \frac{1}{n_k}\right)^{n_k + 1} = \lim_{k \to \infty} \left(1 + \frac{1}{n_k}\right)^{n_k} \lim_{k \to \infty} \left(1 + \frac{1}{n_k}\right) = e,$$

и переходя к пределу в (8.11) при $k \to \infty$, получим

$$\lim_{k \to \infty} (1 + x_k)^{1/x_k} = e. \qquad (8.12)$$

Так как $\{x_k\}$ — произвольная последовательность, удовлетворяющая условиям (8.10), то тем самым доказано, что

$$\lim_{x \to 0} \quad (1 + x)^{1/x} = e. \qquad (8.13)$$

Пусть теперь последовательность $\{x_k\}$ такая, что $\lim\limits_{k \to \infty} x_k = -0$, т. е.

$$\lim_{k \to \infty} x_k = 0, \;\; x_k < 0. \qquad (8.14)$$

Положим $y_k = -x_k$, тогда $y_k > 0$ и $\lim\limits_{k \to \infty} y_k = 0$, причем без ограничения общности можно считать, что $y_k < 1$, $k = 1, 2, \ldots$. Тогда

$$\lim_{k \to \infty} \; (1 + x_k)^{1/x_k} = \lim_{k \to \infty} \; (1 - y_k)^{-1/y_k} = \lim_{k \to \infty} \; \left(\frac{1}{1 - y_k}\right)^{1/y_k} =$$

$$= \lim_{k \to \infty} \; \left(1 + \frac{y_k}{1 - y_k}\right)^{1/y_k} = \lim_{k \to \infty} (1 + z_k)^{1/z_k + 1},$$

где

$$z_k = \frac{y_k}{1 - y_k} > 0 \;\text{ и }\; \lim_{k \to \infty} z_k = 0,$$

и, в силу уже доказанного равенства (8.13),

$$\lim_{k \to \infty} (1 + x_k)^{1/x_k} = \lim_{k \to \infty} (1 + z_k)^{1/z_k} \lim_{k \to \infty} (1 + z_k) = e.$$

Но $\{x_k\}$ — произвольная последовательность, удовлетворяющая условиям (8.14), поэтому

$$\lim_{x \to -0}(1 + x)^{1/x} = e. \tag{8.15}$$

Таким образом, функция $(1 + x)^{1/x}$, $x \neq 0$, имеет в точке 0 пределы слева и справа, равные одному и тому же числу e. Поэтому существует и ее двусторонний предел при $x \to 0$, также равный e (см. п. 5.9).

СЛЕДСТВИЕ 1.

$$\lim_{x \to 0} \frac{\log_a(1 + x)}{x} = \frac{1}{\ln a}, \; a > 0, a \neq 1, \tag{8.16}$$

и, в частности, при $a = e$

$$\lim_{x \to 0} \frac{\ln(1 + x)}{x} = 1.$$

В самом деле, приняв во внимание непрерывность логарифмической функции (см. теорему 4 из § 7), теорему о пределе композиции функций (см. п. 5.16) и равенство (8.6), получим

$$\lim_{x \to 0} \frac{\log_a(1 + x)}{x} = \lim_{x \to 0} \log_a (1 + x)^{1/x} =$$
$$= \log_a \lim_{x \to 0}(1 + x)^{1/x} = \log_a e = \frac{1}{\ln a}.$$

СЛЕДСТВИЕ 2.

$$\lim_{x \to 0} \frac{a^x - 1}{x} = \ln a. \tag{8.17}$$

В частности, при $a = e$ имеет место равенство

$$\lim_{x \to 0} \frac{e^x - 1}{x} = 1. \tag{8.18}$$

Функция $y = a^x - 1$ строго монотонна и непрерывна на всей числовой оси, поэтому обратная функция $x = \frac{\ln(1 + y)}{\ln a}$ также строго монотонна и непрерывна при $y > -1$. При $x = 0$ имеем также и $y = 0$, поэтому обозначения $x \to 0$ и $y \to 0$ эквивалентны (см. замечание 4 в п. 5.16). Воспользуемся для вычисления предела (8.17) правилом замены переменного. Положив $x = \frac{\ln(1 + y)}{\ln a}$, получим

$$\lim_{x \to 0} \frac{a^x - 1}{x} = \lim_{y \to 0} \frac{y \ln a}{\ln(1 + y)} = \ln a \frac{1}{\lim\limits_{y \to 0} \frac{\ln(1 + y)}{y}} = \ln a.$$

8.2. Сравнение функций

Как мы уже знаем, сумма, разность и произведение бесконечно малых функций являются также бесконечно малыми функциями; этого нельзя, вообще говоря, сказать об их частном: деление одной бесконечно малой на другую может привести к разнообразным случаям, как это показывают приведенные ниже примеры бесконечно малых при $x \to 0$ функций $\alpha(x)$ и $\beta(x)$.

Пусть, например, $\alpha(x) = x$ и $\beta(x) = x^2$; тогда

$$\lim_{x \to 0} \frac{\beta(x)}{\alpha(x)} = \lim_{x \to 0} x = 0, \quad \lim_{x \to 0} \frac{\alpha(x)}{\beta(x)} = \lim_{x \to 0} \frac{1}{x} = \infty.$$

Если же $\alpha(x) = x$, $\beta(x) = 2x$, то $\lim\limits_{x \to 0} \frac{\beta(x)}{\alpha(x)} = 2$, а если $\alpha(x) = x$, $\beta(x) = x \sin \frac{1}{x}$, то предел $\lim\limits_{x \to 0} \frac{\beta(x)}{\alpha(x)}$ не существует.

Все рассматриваемые в этом параграфе функции предполагаются определенными на некотором множестве $X \subset \boldsymbol{R}$, под x_0 понимается либо число $x_0 \in \boldsymbol{R}$, либо одна из бесконечностей ∞, $+\infty$ или $-\infty$. В том случае, когда x_0 — число, x_0 является точкой прикосновения множества X, причем содержательным является лишь тот случай, когда x_0 — предельная точка множества X. При этом возможно, что $x_0 \in X$ и что $x_0 \notin X$. Последнее заведомо имеет место, если рассматриваемая функция имеет в точке x_0 какой-либо бесконечный предел. Если же точка x_0 является одной из бесконечностей ∞, $+\infty$, $-\infty$, то множество X предполагается неограниченным, соответственно неограниченным сверху или снизу.

Рассмотрим вопрос сравнения функций в окрестности точки, в частности сравнение бесконечно малых и бесконечно больших функций; эти случаи являются основными.

Определение 1. *Если для функций $f : X \to \boldsymbol{R}$ и $g : X \to \boldsymbol{R}$ существует такая постоянная $c > 0$, что в некоторой окрестности точки x_0 для всех точек $x \in X$ выполняется неравенство*

$$|f(x)| \leqslant c|g(x)|, \tag{8.19}$$

то функцию f называют ограниченной по сравнению с функцией g в окрестности точки x_0 и в этом случае пишут

$$f(x) = O(g(x)), \ x \to x_0$$

(«$f(x)$ есть O большое от $g(x)$, при x, стремящемся к x_0»).

Подчеркнем, что запись $x \to x_0$ имеет здесь другой, чем обычно, смысл: она только указывает на то, что рассматриваемое свойство имеет место лишь в некоторой окрестности точки x_0; ни о каком пределе здесь речи нет.

ЛЕММА 3. *Если* $f(x) = \varphi(x)g(x)$, $x \in X$, *и существует конечный предел*

$$\lim_{x \to x_0} \varphi(x) = k,$$

то

$$f(x) = O(g(x)), \ x \to x_0.$$

Доказательство. Из существования конечного предела $\lim\limits_{x \to x_0} \varphi(x) = k$ (см. свойство 1^0 пределов функций в п. 5.10) следует существование такой окрестности $U(x_0)$ точки x_0, что функция φ ограничена на $X \cap U(x_0)$, т. е. имеется такая постоянная $c > 0$, что для всех $x \in X \cap U(x_0)$ выполняется неравенство $|\varphi(x)| \leqslant c$, следовательно, и неравенство $|f(x)| = |\varphi(x)||g(x)| \leqslant c|g(x)|$. Это, согласно определению 1, и означает, что $f(x) = O(g(x))$, $x \to x_0$. □

Примеры. 1. $\frac{1}{x} = O\left(\frac{1}{x^2}\right)$ при $x \to 0$, поскольку $\left|\frac{1}{x}\right| \leqslant \frac{1}{x^2}$ при $|x| \leqslant 1$.

2. $\frac{1}{x^2} = O\left(\frac{1}{x}\right)$ *при* $x \to \infty$, *так как* $\frac{1}{x^2} \leqslant \left|\frac{1}{x}\right|$ *при* $|x| \geqslant 1$.

Запись

$$f(x) = O(1), \ x \to x_0,$$

означает, что функция f ограничена в некоторой окрестности точки x_0, например $\frac{\operatorname{tg} 2x}{x} = O(1)$ при $x \to 0$, ибо $\lim\limits_{x \to 0} \frac{\operatorname{tg} 2x}{x} = 2$ и, значит, функция $\frac{\operatorname{tg} 2x}{x}$ ограничена в окрестности точки $x = 0$.

Определение 2. *Если функции* $f(x)$ *и* $g(x)$ *такие, что* $f = O(g)$ *и* $g = O(f)$ *при* $x \to x_0$, *то они называются функциями одного порядка при* $x \to x_0$; *это записывается в виде* $f(x) \asymp g(x)$, $x \to x_0$.

Это понятие наиболее содержательно в том случае, когда функции f и g являются либо бесконечно малыми, либо бесконечно большими при $x \to x_0$. Например, функции $\alpha = x$ и $\beta = x\left(2 + \sin \frac{1}{x}\right)$ являются при $x \to 0$ бесконечно малыми одного порядка, поскольку

$$\left|\frac{\alpha}{\beta}\right| = \frac{1}{\left|2 + \sin \frac{1}{x}\right|} \leqslant \frac{1}{2 - \left|\sin \frac{1}{x}\right|} \leqslant 1, \quad \left|\frac{\beta}{\alpha}\right| = \left|2 + \sin \frac{1}{x}\right| \leqslant 2 + \left|\sin \frac{1}{x}\right| \leqslant 3.$$

ЛЕММА 4. *Если существует конечный предел* $\lim\limits_{x \to x_0} \frac{f(x)}{g(x)} = k \neq 0$, *то* $f(x) \asymp g(x)$, $x \to x_0$.

Доказательство. При $x \to x_0$ определен предел дроби $\frac{f(x)}{g(x)}$, поэтому существует такая окрестность $U(x_0)$ точки x_0, что для всех точек $x \in X \cap U(x_0)$ выполняется неравенство $g(x) \neq 0$. Для этих x положим $\varphi(x) = \frac{f(x)}{g(x)}$. Тогда $f(x) = \varphi(x)g(x)$ и $\lim\limits_{x \to x_0} \varphi(x) = k$. Следовательно, по лемме 3, $f(x) = O(g(x))$, $x \to x_0$.

Из условия $\lim\limits_{x \to x_0} \frac{f(x)}{g(x)} \neq 0$ следует, что существует и такая окрестность $U(x_0)$ точки x_0, что для всех $x \in X \cap U(x_0)$ выполняется неравенство $\frac{f(x)}{g(x)} \neq 0$ (см. свойство 2^0 пределов функций в п. 5.10), а следовательно, и неравенство $f(x) \neq 0$. Для $x \in X \cap U(x_0)$ положим $\psi(x) = \frac{g(x)}{f(x)}$; тогда $g(x) = \psi(x)f(x)$ и $\lim\limits_{x \to x_0} \psi(x) = \frac{1}{k}$. Поэтому снова, согласно лемме 3, $g(x) = O(f(x))$, $x \to x_0$. □

В качестве примера возьмем функции $f(x) = 3x^2$ и $g(x) = \sin x^2$. Имеем $\lim\limits_{x \to 0} \frac{g(x)}{f(x)} = \frac{1}{3}\lim \frac{\sin x^2}{x^2} = \frac{1}{3}$ (см. (8.1)), поэтому, согласно лемме 4, функции $3x^2$ и $\sin x^2$ одного порядка при $x \to 0$.

Замечание. Отметим, что условие (8.19) равносильно следующему: существует такая ограниченная функция φ: $X \to \boldsymbol{R}$, что в некоторой окрестности точки x_0 для всех $x \in X$ выполняется равенство $f(x) = \varphi(x)g(x)$.

Действительно, при выполнении этого условия из ограниченности функции φ непосредственно следует неравенство (8.19). Наоборот, если в некоторой окрестности точки x_0 выполнено условие (8.19), то, полагая

$$\varphi(x) = \begin{cases} \frac{f(x)}{g(x)}, \text{ если } g(x) \neq 0, \\ 0, \quad \text{если } g(x) = 0, \end{cases}$$

получаем, что $f(x) = \varphi(x)g(x)$ и функция φ, очевидно, ограничена (для точек $x \in X$ рассматриваемой окрестности точки x_0).

Определение 3. *Функции $f : X \to \boldsymbol{R}$ и $g : X \to \boldsymbol{R}$ называются эквивалентными при $x \to x_0$, если существует такая функция $\varphi : X \to \boldsymbol{R}$, что в некоторой окрестности точки x_0 для всех точек $x \in X$ выполняется равенство*

$$f = \varphi(x)g(x) \tag{8.20}$$

и

$$\lim_{x \to x_0} \varphi(x) = 1. \tag{8.21}$$

Сразу заметим, что существование предела у функции в данной точке является локальным свойством функции, поэтому значения функции φ вне указанной в определении окрестности не играют роли.

Если выполнено свойство (8.21), то найдется такая окрестность $U = U(x_0)$ точки x_0, что при $x \in X \cap U$ выполняется неравенство $\varphi(x) \neq 0$ (см. свойство 2^0 пределов функций в п. 5.10). Полагая $\psi(x) = \frac{1}{\varphi(x)}$, $x \in X \cap U$, видим, что условия (8.20) и (8.21) равносильны условиям

$$g(x) = \psi(x)f(x), \;\; x \in X \cap U, \tag{8.20$'$}$$

$$\lim_{x \to x_0} \psi(x) = 1. \tag{8.21$'$}$$

Таким образом, если функции f и g эквивалентны при $x \to x_0$, то и функции g и f также эквивалентны при $x \to x_0$, т. е., как говорят, эквивалентность двух функций обладает свойством симметричности.

Отметим, что свойство функций быть функциями одного порядка также является симметричным свойством, а свой-

ство одной функции быть «O большим» относительно другой уже не симметрично.

Функции $f(x)$ и $g(x)$, эквивалентные при $x \to x_0$, называются также *асимптотически равными* при $x \to x_0$. *Асимптотическое равенство* (эквивалентность) функций обозначается символом $\sim$:

$$f(x) \sim g(x) \text{ при } x \to x_0. \qquad (8.22)$$

Из сказанного выше следует, что если $f \sim g$ при $x \to x_0$, то и $g \sim f$ при $x \to x_0$.

Отметим, что если в условиях определения 3 функции f и g, а следовательно, и функция φ, определены в точке x_0, т. е. $x_0 \in X$, то предел (8.21) берется при $x \to x_0$ по множеству, содержащему точку x_0, поэтому функция φ непрерывна в точке x_0; тогда из условия (8.21) следует, что

$$\varphi(x_0) = 1.$$

Примеры. 1. $\dfrac{x^2}{1+x^4} \sim x^2$ при $x \to 0$. Действительно, полагая $\varphi(x) = \dfrac{1}{1+x^4}$, получаем

$$\frac{x^2}{1+x^4} = \varphi(x)x^2 \text{ и } \lim_{x\to 0} \varphi(x) = \lim_{x\to 0} \frac{1}{1+x^4} = 1.$$

2. $\dfrac{x^6}{1+x^4} \sim x^2$ при $x \to \infty$. В самом деле, если $\varphi(x) = \dfrac{x^4}{1+x^4}$, то

$$\frac{x^6}{1+x^4} = \varphi(x)x^2 \text{ и } \lim_{x\to\infty} \varphi(x) = \lim_{x\to\infty} \frac{x^4}{1+x^4} = 1.$$

Пусть существует такая проколотая окрестность $\mathring{U} = \mathring{U}(x_0)$ точки x_0, что для всех $x \in X \cap \mathring{U}$ выполняются неравенства $f(x) \neq 0$ и $g(x) \neq 0$ и $x_0 \in X$. Тогда условия (8.20) и (8.21) равносильны соотношению $\lim\limits_{\substack{x\to x_0 \\ x \in X\cap\mathring{U}}} \dfrac{f(x)}{g(x)} = 1$ и, следовательно, соотношению $\lim\limits_{\substack{x\to x_0 \\ x \in X\cap\mathring{U}}} \dfrac{g(x)}{f(x)} = 1.$

Действительно, ясно, что эти соотношения при сделанных предположениях о необращении в нуль функций f и g на множестве $X \cap \overset{\circ}{U}$ и $x_0 \notin X$ сразу вытекают из условий (8.20) и (8.21). Наоборот, если они выполнены, то достаточно положить $\varphi(x) = \frac{f(x)}{g(x)}$, $x \in X \cap \overset{\circ}{U}(x_0)$ и $\varphi(x_0) = 1$, тогда, очевидно, для функции φ выполняются условия (8.20) и (8.21).

Если

$$f \sim g \text{ и } g \sim h \text{ при } x \to x_0, \tag{8.23}$$

то

$$f \sim h \text{ при } x \to x_0. \tag{8.24}$$

В самом деле, из условий (8.23) следует, что существуют такие окрестность $U = U(x_0)$ точки x_0 и функции φ: $X \cap U \to \to \boldsymbol{R}$ и ψ: $X \cap U \to \boldsymbol{R}$, что для всех $x \in X \cap U$ имеют место равенства

$$f(x) = \varphi(x)g(x), \quad g(x) = \psi(x)h(x)$$

и

$$\lim_{x \to x_0} \varphi(x) = \lim_{x \to x_0} \psi(x) = 1,$$

поэтому

$$f(x) = \varphi(x)\psi(x)h(x),$$

где $\lim\limits_{x \to x_0} \varphi(x)\psi(x) = 1$, т. е. выполняется асимптотическое равенство (8.24).

Из результатов п. 8.1 следует, что при $x \to 0$ справедлива следующая эквивалентность бесконечно малых:

$$x \sim \sin x \sim \operatorname{tg} x \sim \arcsin x \sim \operatorname{arctg} x \sim \ln(1 + x) \sim e^x - 1.$$

Из этой эквивалентности следуют и более общие соотношения, которые сформулируем в виде отдельной леммы.

ЛЕММА 5. *Если функция $u(x)$ такова, что*

$$\lim_{x \to x_0} u(x) = 0, \tag{8.25}$$

то при $x \to x_0$

$$u(x) \sim \sin u(x) \sim \operatorname{tg} u(x) \sim \arcsin u(x) \sim$$
$$\sim \operatorname{arctg} u(x) \sim \ln[1 + u(x)] \sim e^{u(x)} - 1. \tag{8.26}$$

Доказательство. Покажем, например, что

$$\sin u(x) \sim u(x) \text{ при } x \to x_0, \tag{8.27}$$

где u: $X \to \boldsymbol{R}$ и $\lim\limits_{x \to x_0} u(x) = 0$. Определим для всех $x \in X$ функцию φ: $X \to \boldsymbol{R}$ следующим образом:

$$\varphi(x) = \begin{cases} \dfrac{\sin u(x)}{u(x)}, & \text{если } u(x) \neq 0, \\ 1, & \text{если } u(x) = 0, \end{cases} \tag{8.28}$$

и покажем, что

$$\lim_{x \to x_0} \varphi(x) = 1. \tag{8.29}$$

Для этого разобьем множество X на два подмножества:

$$X_1 = \{x \in X : u(x) \neq 0\} \text{ и } X_2 = \{x \in X : u(x) = 0\}. \tag{8.30}$$

Пусть сначала множества X_1 и X_2 не пусты, а x_0 является конечной или бесконечно удаленной точкой прикосновения каждого из них. Функция $\dfrac{\sin u(x)}{u(x)}$ определена на множестве X_1 и по теореме о пределе сложной функции (см. теорему 6 в п. 5.16) имеем

$$\lim_{\substack{x \to x_0 \\ x \in X_1}} \varphi(x) = \lim_{\substack{x \to x_0 \\ x \in X_1}} \frac{\sin u(x)}{u(x)} = \lim_{u \to 0} \frac{\sin u}{u} = 1.$$

Здесь было использовано еще одно свойство пределов функций: если функция $\left(\text{в данном случае } \dfrac{\sin u}{u}\right)$ имеет предел при $u \to u_0$ по некоторому множеству, то она, согласно лемме 1 п. 5.4, имеет тот же предел при $u \to u_0$ по любому подмножеству этого множества (в рассматриваемом случае по подмножеству числовой оси, состоящему из множества значений функции $u(x)$, отличных от нуля).

На множестве же X_2 функция φ тождественно равна 1, поэтому

$$\lim_{\substack{x \to x_0 \\ x \in X_2}} \varphi(x) = \lim_{\substack{x \to x_0 \\ x \in X_2}} 1 = 1.$$

Таким образом, на каждом из множеств X_1 и X_2 функция φ при $x \to x_0$ имеет один и тот же предел, равный 1, а так как $X = X_1 \cup X_2$, то тот же предел при $x \to x_0$ она имеет и по всему множеству (см. лемму 5 в п. 5.8), т. е. в рассмотренном случае равенство (8.29) доказано.

Если же одно из множеств X_1 или X_2 окажется пустым или точка x_0 не является точкой прикосновения (конечной или бесконечно удаленной) одного из них, то равенство (8.29) также будет иметь место, так как в этих случаях предел функции φ при $x \to x_0$ по множеству X сведется к пределу по одному из множеств X_1 или X_2, для которых, как уже установлено, рассматриваемый предел равен единице.

Итак, равенство (8.29) доказано, а так как из (8.28) следует, что для всех $x \in X \cap U(x_0)$ имеет место соотношение $\sin u(x) = \varphi(x)u(x)$, то доказана и справедливость асимптотического равенства (8.27).

Аналогично доказываются и остальные асимптотические формулы (8.26). □

Определение 4. *Функция* α: $X \to \boldsymbol{R}$ *называется бесконечно малой при* $x \to x_0$ *по сравнению с функцией* f: $X \to \boldsymbol{R}$, *если существует такая функция* ε: $X \to \boldsymbol{R}$, *что в некоторой окрестности точки* x_0 *для всех* $x \in X$ *выполняется равенство*

$$\alpha(x) = \varepsilon(x)f(x) \tag{8.31}$$

и

$$\lim_{x \to x_0} \varepsilon(x) = 0. \tag{8.32}$$

Как и для эквивалентных функций, в силу локальности свойства существования предела функции в точке, значения функции $\varepsilon(x)$ вне указанной в определении окрестности несущественны и в случае необходимости могут быть выбраны произвольно.

Если функция α является бесконечно малой при $x \to x_0$ по сравнению с функцией f, то пишут

$$\alpha(x) = o(f(x)), \quad x \to x_0$$

(читается: «$\alpha(x)$ есть *o* малое от $f(x)$ при $x \to x_0$»).

В силу этого определения, например, запись «$\alpha(x) = o(1)$, $x \to x_0$» означает просто, что функция α является бесконечно малой при $x \to x_0$.

Отметим, что если в условии определения 4 бесконечно малой функции по сравнению с другой функции α и f, а сле-

довательно, и функция $\varepsilon(x)$ определены в точке x_0, т. е. $x_0 \in \in X$, то предел (8.32) берется при $x \to x_0$ по множеству, содержащему точку x_0. Поэтому в этом случае функция $\varepsilon(x)$ является непрерывной в точке x_0 и, следовательно, согласно (8.33), имеет место равенство $\varepsilon(x_0) = 0$. Этот факт мы неоднократно будем использовать в дальнейшем.

Если существует такая проколотая окрестность $\mathring{U} = \mathring{U}(x_0)$ точки x_0, что для всех точек $x \in X \cap \mathring{U}$ выполняется неравенство $f(x) \neq 0$, а в случае $x_0 \in X$, кроме того, $\alpha(x_0) = 0$, то условия (8.31)—(8.32) равносильны условию

$$\lim_{\substack{x \to x_0 \\ x \in X \cap \mathring{U}}} \frac{\alpha(x)}{f(x)} = 0. \tag{8.33}$$

В самом деле, при сделанном предположении о неравенстве нулю функции f условие (8.33) сразу следует из (8.31)—(8.32). Наоборот, если выполнено (8.33), то достаточно положить

$$\varepsilon(x) = \frac{\alpha(x)}{f(x)}, \; x \in X \cap \mathring{U},$$

а если $x_0 \in X$, то еще $\varepsilon(x_0) = 0$, чтобы были выполнены условия (8.31)—(8.32).

В том случае, когда $f(x)$ сама является бесконечно малой при $x \to x_0$, говорят, что функция $\alpha = o(f)$ при $x \to x_0$ есть бесконечно малая более высокого порядка, чем f. При этом если $\alpha \asymp (f(x))^n$), $x \to x_0$, то бесконечно малая α называется бесконечно малой порядка n относительно бесконечно малой f, $n = 1, 2, \ldots$.

Например, $x^3 = o(\sin x^2)$ при $x \to 0$, так как

$$\lim_{x \to 0} \frac{x^3}{\sin x^2} = \lim_{x \to 0} x \lim \frac{x^2}{\sin x^2} = 0 \cdot 1 = 0.$$

Бесконечно малая $\alpha(x) = x^3$ является бесконечно малой более высокого порядка, чем бесконечно малая $f(x) = = \sin x^2$, $x \to 0$.

Бесконечно малая $\alpha(x) = 1/x^4$ является бесконечно малой четвертого порядка относительно бесконечно малой $f(x) = 1/x$ при $x \to \infty$.

Конечно, символ «o малое» можно применять не только к бесконечно малым: например, $x = o(x^2)$ при $x \to \infty$.

Отметим, что если $f = o(g)$ при $x \to x_0$, то и подавно $f = O(g)$ при $x \to x_0$. В самом деле, пусть $f = \varepsilon g$, где $\lim\limits_{x \to x_0} \varepsilon = 0$. Тогда функция $\varepsilon = \varepsilon(x)$ ограничена на пересечении множества X с некоторой окрестностью $U(x_0)$ точки x_0 (см. п. 5.10): $|\varepsilon(x)| \leqslant c$ и, следовательно, $|f(x)| \leqslant c|g(x)|$ для всех $x \in X \cap U(x_0)$. А это и означает, что $f = O(g)$, $x \to x_0$.

Соберем вместе введенные в этом пункте основные понятия. Пусть заданы функции $f\colon X \to \boldsymbol{R}$ и $g\colon X \to \boldsymbol{R}$ и существует такая функция $\varphi\colon X \to \boldsymbol{R}$, что для всех точек множества X, лежащих в некоторой окрестности $U = U(x_0)$ точки x_0, выполняется равенство $f(x) = \varphi(x)g(x)$, тогда:

если функция $\varphi(x)$ ограничена на U, то $f(x) = O(g(x))$, $x \to x_0$;

если $\lim\limits_{x \to x_0} \varphi(x) = 1$, то $f(x) \sim g(x)$, $x \to x_0$;

если $\lim\limits_{x \to x_0} \varphi(x) = 0$, то $f(x) = o(g(x))$, $x \to x_0$.

УПРАЖНЕНИЕ. Пусть $\beta = O(\alpha^2)$ при $x \to x_0$, $\lim\limits_{x \to x_0} \alpha = 0$. Доказать, что в этом случае $\beta = o(\alpha)$ при $x \to x_0$.

Отметим, что в частном случае множество X может быть множеством натуральных чисел $\boldsymbol{N}$, и тогда при $x_0 = +\infty$ мы получим понятия последовательности $\{x_n\}$, ограниченной по сравнению с последовательностью $\{y_n\}$:

$$x_n = O(y_n), \quad n \to \infty,$$

последовательности $\{x_n\}$ одного порядка с последовательностью $\{y_n\}$

$$x_n \asymp y_n, \quad n \to \infty,$$

последовательности $\{x_n\}$, асимптотически равной (эквивалентной) последовательности $\{y_n\}$,

$$x_n \sim y_n, \quad n \to \infty,$$

и последовательности $\{x_n\}$, бесконечно малой по сравнению с последовательностью $\{y_n\}$,

$$x_n = o(y_n), \quad n \to \infty.$$

При использовании равенств с символами O и o следует иметь в виду, что они не являются равенствами в обычном смысле этого слова. Так, если $\alpha_1 = o(\beta)$ при $x \to x_0$, $\alpha_2 = o(\beta)$ при $x \to x_0$, то было бы ошибкой сделать отсюда заключение, что $\alpha_1 = \alpha_2$, как это было бы в случае обычных равенств. Например, $x^3 = o(x)$ и $x^2 = o(x)$ при $x \to 0$, но $x^2 \neq x^3$.

Аналогично, если $f + O(f) = g + O(f)$ при $x \to x_0$, то было бы ошибкой сделать заключение, что $f = g$.

Дело в том, что один и тот же символ $O(f)$ или $o(f)$ может обозначать разные конкретные функции. Это обстоятельство связано с тем, что при определении символов $O(f)$ и $o(f)$ мы по существу ввели целые классы функций, обладающих определенными свойствами (класс функций, ограниченных в некоторой окрестности точки x_0 по сравнению с функцией f, и класс функций, бесконечно малых по сравнению с $f(x)$ при $x \to x_0$), и было бы правильнее писать не $\alpha = O(f)$ и $\alpha = o(f)$, а соответственно $\alpha \in O(f)$ и $\alpha \in o(f)$. Однако это привело бы к существенному усложнению вычислений по формулам, в которых встречаются символы O и o. Поэтому мы сохраним прежнюю запись $\alpha = O(f)$ и $\alpha = o(f)$, но будем всегда читать эти равенства, в соответствии с приведенными выше определениями, только в одну сторону: слева направо (если, конечно, не оговорено что-либо другое). Например, запись:

$$\alpha = o(f), \ x \to x_0,$$

означает, что функция α является бесконечно малой по сравнению с функцией f при $x \to x_0$, но отнюдь не то, что всякая бесконечно малая по сравнению с f функция равна α.

В качестве примера использования этих символов докажем равенство

$$o(cf) = o(f), \tag{8.34}$$

где c — постоянная.

Согласно сказанному, надо показать, что если $g = o(cf)$, то $g = o(f)$. Действительно, если $g = o(cf)$, то $g = \varepsilon cf$, где $\lim\limits_{x \to x_0} \varepsilon(x) = 0$. Положим $\varepsilon_1 = c\varepsilon$; тогда $g = \varepsilon_1 f$, где, очевидно, $\lim\limits_{x \to x_0} \varepsilon_1(x) = 0$ и, значит, $g = o(f)$. □

В заключение отметим, что сказанное об использовании символов o и O не исключает, конечно, того, что отдельные

формулы с этими символами могут оказаться справедливыми не только при чтении слева направо, но и справа налево; так, формула (8.34) при $c \neq 0$ верна и при чтении справа налево.

8.3. Эквивалентные функции

Если функция $f(x)$ заменяется функцией $g(x)$, то разность $f(x) - g(x)$ называется *абсолютной погрешностью*, а отношение $\frac{f(x) - g(x)}{f(x)}$ — *относительной погрешностью* сделанной замены. Если изучается поведение функции $f(x)$ при $x \to x_0$, то часто целесообразно заменить ее функцией $g(x)$ такой, что: 1) функция $g(x)$ в определенном смысле более простая, чем функция $f(x)$; 2) абсолютная погрешность стремится к нулю при $x \to x_0$:

$$\lim_{x \to x_0} [f(x) - g(x)] = 0.$$

В этом случае говорят, что $g(x)$ приближает или аппроксимирует функцию $f(x)$ вблизи точки x_0. Таким свойством обладают, например, все бесконечно малые при $x \to x_0$ функции f и g.

Ниже будет показано, что среди них лишь те, которые эквивалентны между собой: $g(x) \sim f(x)$, $x \to x_0$, обладают тем свойством, что не только абсолютная погрешность $f(x) - g(x)$, но и относительная $\frac{f(x) - g(x)}{f(x)}$ стремится к нулю при $x \to x_0$:

$$\lim_{x \to x_0} \frac{f(x) - g(x)}{f(x)} = 0.$$

В этом смысле функции, эквивалентные заданной, приближают ее лучше, чем другие функции.

Например, функции x, $\frac{1}{2}x$, $2x$, $10x$ являются бесконечно малыми при $x \to 0$, так же как и $\sin x$, а поэтому абсолютные погрешности при замене $\sin x$ каждой из них стремятся к нулю при $x \to 0$:

$$\lim_{x \to 0} (\sin x - x) = \lim_{x \to 0} (\sin x - \tfrac{1}{2}x) =$$
$$= \lim_{x \to 0} (\sin x - 2x) = \lim_{x \to 0} (\sin x - 10x) = 0.$$

Но лишь одна из всех перечисленных функций, а именно $g(x) = x$, обладает тем свойством, что относительная погреш-

ность при замене $\sin x$ этой функцией будет стремиться к нулю при $x \to 0$:

$$\lim_{x \to 0} \frac{\sin x - x}{\sin x} = \lim_{x \to 0} \left(1 - \frac{x}{\sin x}\right) = 0.$$

Стремление относительной погрешности $\frac{f(x) - g(x)}{f(x)}$ к нулю при $x \to x_0$ можно записать, используя символ «o малое»:

$$f(x) - g(x) = o(f(x)),\ x \to x_0.$$

Сформулируем высказанное характеристическое свойство эквивалентных функций в виде теоремы.

ТЕОРЕМА 1. *Для того чтобы функции f: $X \to \boldsymbol{R}$ и g: $X \to \boldsymbol{R}$ были эквивалентны при $x \to x_0$, необходимо и достаточно, чтобы при $x \to x_0$ выполнялось условие*

$$f(x) = g(x) = o(g(x)),\ x \to x_0. \tag{8.35}$$

Доказательство. Формула (8.35) является просто другой записью определения 3. Действительно, условие (8.21): $\lim\limits_{x \to x_0} \varphi(x) = 1$ равносильно условию $\varphi(x) = 1 + \varepsilon(x)$, где $\lim\limits_{x \to x_0} \varepsilon(x) = 0$. Поэтому условие

$$f(x) = \varphi(x)g(x),\ \lim_{x \to x_0} \varphi(x) = 1,$$

равносильно условию

$$f(x) = (1 + \varepsilon(x))g(x) = g(x) + \varepsilon(x)g(x),\ \lim_{x \to x_0} \varepsilon(x) = 0,$$

т. е. условию $f(x) = g(x) + o(g(x))$, $x \to x_0$. □

Пример 1. $\operatorname{ctg} x = \frac{1}{x} + o\left(\frac{1}{x}\right)$, $x \to 0$.

Действительно, в силу теоремы 1, достаточно показать, что $\operatorname{ctg} x \sim \frac{1}{x}$, $x \to 0$. Это же сразу следует из равенства (8.3):

$$\lim_{x \to 0} \frac{\operatorname{ctg} x}{1/x} = \lim_{x \to 0} \frac{x}{\operatorname{tg} x} = 1.$$

В том случае, когда существует такая проколотая окрестность $\overset{\circ}{U}(x_0)$ точки x_0, что функции f: $X \to \boldsymbol{R}$, g: $X \to \boldsymbol{R}$ не обращаются в нуль на пересечении $X \cap \overset{\circ}{U}(x_0)$ и $x_0 \notin X$, теорема 1, очевидно, равносильна утверждению, что функции f и g эквивалентны при $x \to x_0$ тогда и только тогда, ког-

да относительная погрешность $\frac{f(x)-g(x)}{f(x)}$ $\Bigl($или, в силу симметричности понятия эквивалентности функций, отношение $\frac{f(x)-g(x)}{g(x)}\Bigr)$ стремится к нулю при $x \to x_0$.

СЛЕДСТВИЕ. *Пусть* $\lim\limits_{x \to x_0} \frac{g(x)}{f(x)} = c \neq 0$. *Тогда* $g \sim cf$ *и* $g(x) = cf(x) + o(f(x))$ *при* $x \to x_0$.

Доказательство. Если $\lim\limits_{x \to x_0} \frac{g(x)}{f(x)} = c \neq 0$, то $\lim\limits_{x \to x_0} \frac{g(x)}{cf(x)} = 1$, и поэтому $g \sim cf$ при $x \to x_0$. Отсюда по теореме 1 имеем $g(x) = cf(x) + o(cf(x))$, откуда (см. конец п. 8.2) $g(x) = cf(x) + o(f(x))$, $x \to x_0$. □

ТЕОРЕМА 2. *Пусть функции* f, f_1, g, g_1 *заданы на множестве* X *и* $f(x) \sim f_1(x)$, $g(x) \sim g_1(x)$ *при* $x \to x_0$. *Тогда если существует*

$$\lim_{x \to x_0} \frac{f_1(x)}{g_1(x)}, \tag{8.36}$$

то существует и $\lim\limits_{x \to x_0} \frac{f(x)}{g(x)}$, *причем*

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{f_1(x)}{g_1(x)}. \tag{8.37}$$

Доказательство. Условия $f \sim f_1$ и $g \sim g_1$, $x \to x_0$, $x \in X$, означают, что существуют такая окрестность $U = U(x_0)$ точки x_0 и такие функции φ и ψ, определенные на пересечении $X \cap U$, что

$$f(x) = \varphi(x) f_1(x), \quad g(x) = \psi(x) g_1(x), \quad x \in X \cap U,$$

$$\lim_{x \to x_0} \varphi(x) = \lim_{x \to x_0} \psi(x) = 1.$$

Поэтому

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{\varphi(x) f_1(x)}{\psi(x) g_1(x)} =$$

$$= \frac{\lim\limits_{x \to x_0} \varphi(x)}{\lim\limits_{x \to x_0} \psi(x)} \lim_{x \to x_0} \frac{f_1(x)}{g_1(x)} = \lim_{x \to x_0} \frac{f_1(x)}{g_1(x)},$$

т. е. имеет место равенство (8.37). К этому можно лишь добавить, что из существования предела $\lim\limits_{x \to x_0} \psi(x) = 1$ следует, что окрестность $U = U(x_0)$ можно выбрать таким образом, что для всех точек $x \in X \cap U$ будет выполняться неравенство $\psi(x) \neq 0$, а из существования по множеству X предела $\lim\limits_{x \to x_0} \frac{f(x)}{g(x)}$ следует, что окрестность U может быть выбрана еще и так, что для всех $x \in X \cap U$ будет выполняться неравенство $g(x) \neq 0$, так как частное $\frac{f(x)}{g(x)}$ должно быть определено на пересечении $X \cap U$ множества X с некоторой окрестностью U точки x_0. Поэтому все написанные выше выражения имеют смысл. □

Обе части равенства (8.37) равноправны, поэтому из доказанной теоремы следует, что предел, стоящий в левой части, существует тогда и только тогда, когда существует предел в правой части, причем в случае их существования они совпадают. Это делает очень удобным применение теоремы 2 на практике: ее можно использовать для вычисления пределов, не зная заранее, существует или нет рассматриваемый предел.

8.4. Метод выделения главной части функции и его применение к вычислению пределов

Пусть заданы функции $\alpha: X \to \boldsymbol{R}$ и $\beta: X \to \boldsymbol{R}$. Если функция β для всех $x \in X$ представима в виде

$$\beta(x) = \alpha(x) + o(\alpha(x)),\ x \to x_0,$$

то функция α называется *главной частью* функции β при $x \to x_0$.

Примеры. 1. Главная часть функции $\sin x$ при $x \to 0$ равна x, ибо $\sin x = x + o(x)$ при $x \to 0$.

2. Если $P_n(x) = a_n x^n + \ldots + a_1 x + a_0$, $a_n \neq 0$, то функция $a_n x^n$ является главной частью многочлена $P_n(x)$ при $x \to \infty$, так как $P_n(x) = a_n x^n + o(x^n)$ при $x \to \infty$.

Если задана функция $\beta: X \to \boldsymbol{R}$, то ее главная часть при $x \to x_0$ не определяется однозначно: согласно теореме 1, любая функция α, эквивалентная β при $x \to x_0$, является ее главной частью при $x \to x_0$.

Например, пусть $\beta = x + x^2 + x^3$. Так как, с одной стороны, $x^2 + x^3 = o(x)$ при $x \to 0$, то $\beta = x + o(x)$ при $x \to 0$, а с другой стороны, $x^3 = o(x + x^2)$ при $x \to 0$, поэтому

$$\beta = x + x^2 + o(x + x^2) \text{ при } x \to 0.$$

В первом случае главной частью можно считать $\alpha = x$, во втором $\alpha = x + x^2$. Однако если задаваться определенным видом главной части, то при его разумном выборе можно добиться того, что главная часть указанного вида будет определена однозначно.

В частности, справедлива следующая лемма.

ЛЕММА 5. *Пусть $X \subset \boldsymbol{R}$, $x_0 \in \boldsymbol{R}$ и x_0 — предельная точка множества X. Если функция $\beta : X \to \boldsymbol{R}$ обладает при $x \to x_0$ главной частью вида $A(x - x_0)^k$, $A \neq 0$, где A и k — постоянные, то среди всех главных частей такого вида она определяется единственным образом.*

Действительно, пусть при $x \to x_0$

$$\beta(x) = A(x - x_0)^k + o((x - x_0)^k),\ A \neq 0,$$

и

$$\beta(x) = A_1(x - x_0)^{k_1} + o((x - x_0)^{k_1}),\ A_1 \neq 0.$$

Тогда $\beta(x) \sim A(x - x_0)^k$ и $\beta(x) \sim A_1(x - x_0)^{k_1}$ при $x \to x_0$, $x \in X$. Поэтому $A(x - x_0)^k \sim A_1(x - x_0)^{k_1}$, $x \to x_0$, $x \in X$, т. е.

$$1 = \lim_{x \to x_0} \frac{A(x - x_0)^k}{A_1(x - x_0)^{k_1}} = \frac{A}{A_1} \lim_{x \to x_0} (x - x_0)^{k - k_1},$$

что справедливо лишь в случае $A = A_1$ и $k = k_1$. □

Понятие главной части функции полезно при изучении бесконечно малых и бесконечно больших и с успехом используется при решении разнообразных задач математического анализа. Довольно часто удается бесконечно малую

сложного аналитического вида заменить в окрестности данной точки с точностью до бесконечно малых более высокого порядка более простой (в каком-то смысле) функцией. Например, если $\beta(x)$ удается представить в виде $\beta(x) = A(x - x_0)^k + o((x - x_0)^k)$, то это означает, что с точностью до бесконечно малых более высокого порядка, чем $(x - x_0)^k$ при $x \to x_0$, бесконечно малая $\beta(x)$ ведет себя в окрестности точки x как степенна́я функция $A(x - x_0)^k$.

Покажем на примерах, как метод выделения главной части бесконечно малых применяется к вычислению пределов функций. При этом будем широко использовать полученные соотношения эквивалентности (8.26).

Пусть требуется найти предел (а значит, и доказать, что он существует)

$$\lim_{x \to 0} \frac{\ln(1 + x + x^2) + \arcsin 3x - 5x^3}{\sin 2x + \operatorname{tg}^2 x + (e^x - 1)^5}.$$

Используя доказанную выше (см. соотношения (8.26)) эквивалентность $\ln(1 + u) \sim u$ при $u \to 0$, имеем $\ln(1 + x + x^2) \sim x + x^2$ при $x \to 0$, поэтому (см. теорему 1) $\ln(1 + x + x^2) = x + x^2 + o(x + x^2)$. Однако $o(x + x^2) = o(x)$ (почему?) и $x^2 = o(x)$ при $x \to 0$, следовательно,

$$\ln(1 + x + x^2) = x + o(x) \text{ при } x \to 0.$$

Далее, $\arcsin 3x \sim 3x$, поэтому

$$\arcsin 3x = 3x + o(3x) = 3x + o(x).$$

Очевидно также, что $5x^3 = o(x)$. Из асимптотического равенства $\sin 2x \sim 2x$ получаем

$$\sin 2x = 2x + o(2x) = 2x + o(x),$$

из $\operatorname{tg}^2 x \sim x^2$ будем иметь

$$\operatorname{tg}^2 x = x^2 + o(x^2) = o(x),$$

а из $(e^x - 1)^5 \sim x^5$, аналогично,

$$(e^x - 1)^5 = x^5 + o(x^5) = o(x).$$

Все эти соотношения выполняются при $x \to 0$. Теперь имеем

$$\ln(1 + x + x^2) + \arcsin 3x - 5x^3 =$$
$$= x + o(x) + 3x + o(x) - o(x) = 4x + o(x),$$

$$\sin 2x + \operatorname{tg}^2 x + (e^x - 1)^5 = 2x + o(x) + o(x) = 2x + o(x),$$

поэтому

$$\lim_{x \to 0} \frac{\ln(1 + x + x^2) + \arcsin 3x - 5x^3}{\sin 2x + \operatorname{tg}^2 x + (e^x - 1)^5} = \lim_{x \to 0} \frac{4x + o(x)}{2x + o(x)}.$$

Но $4x + o(x) \sim 4x$, а $2x + o(x) \sim 2x$ при $x \to 0$ и, значит, по теореме 2,

$$\lim_{x \to 0} \frac{4x + o(x)}{2x + o(x)} = \lim_{x \to 0} \frac{4x}{2x} = 2.$$

Таким образом, искомый предел существует и равен 2.

При вычислении пределов функций с помощью метода выделения главной части следует иметь в виду, что в случаях, не рассмотренных в п. 8.3, вообще говоря, нельзя бесконечно малые заменять им эквивалентными. Так, например, при отыскании предела выражения $\lim\limits_{x \to 0} \frac{\sin x - x}{x^3}$ было бы ошибкой заменить функцию $\sin x$ эквивалентной ей при $x \to 0$ функцией x. Естественный метод решения подобных задач будет дан в п. 13.4.

Для отыскания пределов выражений вида $u(x)^{v(x)}$ целесообразно находить предел их логарифмов. Рассмотрим подобный пример. Найдем предел $\lim\limits_{x \to 0} \cos^{1/x^2} 2x$. Из равенства

$$\cos^{1/x^2} 2x = e^{\ln \cos^{1/x^2} 2x} \tag{8.38}$$

видно, что достаточно вычислить предел

$$\lim_{x \to 0} \ln \cos^{1/x^2} 2x = \lim_{x \to 0} \frac{\ln \cos 2x}{x^2} = \frac{1}{2} \lim_{x \to 0} \frac{\ln(1 - \sin^2 2x)}{x^2}.$$

А так как

$$\ln(1 - \sin^2 2x) \sim -\sin^2 2x \sim -(2x)^2 = -4x^2,$$

то отсюда, согласно теореме 2 этого параграфа, имеем

$$\frac{1}{2}\lim_{x\to 0}\frac{\ln(1-\sin^2 2x)}{x^2} = -\frac{1}{2}\lim_{x\to 0}\frac{4x^2}{x^2} = -2;$$

таким образом,

$$\lim_{x\to 0}\ln\cos^{1/x^2} 2x = -2.$$

В силу непрерывности показательной функции, из (8.38) имеем

$$\lim_{x\to 0}\cos^{1/x^2} 2x = e^{\lim\limits_{x\to 0}\ln\cos^{1/x^2} 2x} = 1/e^2.$$

Способ вычисления пределов с помощью выделения главной части функции является очень удобным, простым и вместе с тем весьма общим методом. Некоторое затруднение в его применении связано пока с тем, что еще нет достаточно общего способа выделения главной части функции. Это затруднение будет устранено в дальнейшем (см. § 13).

§ 9.

Производная и дифференциал

9.1. Определение производной

Определение 1. *Пусть функция $y = f(x)$ определена в некоторой окрестности точки $x_0 \in \boldsymbol{R}$ и пусть x — произвольная точка этой окрестности. Если отношение $\frac{f(x)-f(x_0)}{x-x_0}$ имеет предел при $x \to x_0$, то этот предел называется производной функции f в точке x_0 или, что то же, при $x = x_0$, и обозначается $f'(x_0)$:*

$$f'(x_0) = \lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}. \tag{9.1}$$

Если ввести обозначение $x - x_0 = \Delta x$, то определение (9.1) запишется в виде

$$f'(x_0) = \lim_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}.$$

Полагая $f(x_0 + \Delta x) - f(x_0) = \Delta y$, опуская обозначения аргумента и обозначая производную просто через y', получаем еще одну запись определения производной:

$$y' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

Если для некоторого значения x_0 существуют пределы $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \infty$, или $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = +\infty$, или $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = -\infty$, то говорят, что при $x = x_0$ существует *бесконечная производная* или соответственно *бесконечная производная определенного знака*, равная $+\infty$ или $-\infty$.

В дальнейшем под выражением «функция имеет производную» будем понимать всегда наличие конечной производной, если не оговорено противное.

Введем для удобства следующее определение. Пересечение окрестности точки $x_0 \in \boldsymbol{R}$ с лучом $x \geqslant x_0$ ($x \leqslant x_0$) назовем *правосторонней* (*левосторонней*) *окрестностью* точки x_0.

Определение 2. *Если функция f определена в некоторой правосторонней* (*левосторонней*) *окрестности точки x_0 и существует конечный или бесконечный* (*определенного знака*) *предел*

$$\lim_{\Delta x \to +0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \quad \left(\lim_{\Delta x \to -0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}\right),$$

то он называется соответственно конечной или бесконечной правой (*левой*) *производной функции f в точке x_0 и обозначается $f'_+(x_0)$* (*или $f'_-(x_0)$*).

Правая и левая производные называются также правосторонней, соответственно левосторонней, а и та и другая — односторонними производными.

Из теоремы об односторонних пределах (см. п. 5.9) следует, что функция $f(x)$, определенная в некоторой окрестности точки x_0, имеет производную $f'(x_0)$ тогда и только тогда, когда $f'_-(x_0)$ и $f'_+(x_0)$ существуют и $f'_-(x_0) = f'_+(x_0)$. В этом случае $f'(x_0) = f'_-(x_0) = f'_+(x_0)$.

Если функция $f(x)$ определена на некотором промежутке и в каждой его точке существует производная (причем под

производной в конце этого промежутка, который принадлежит промежутку, естественно, понимается соответствующая односторонняя производная), то она, очевидно, также является функцией, определенной на данном промежутке; ее обозначают через $f'(x)$. Если $y = f(x)$, то вместо $f'(x_0)$ пишут также $y'|_{x = x_0}$.

Вычисление производной от функции называется *дифференцированием*.

Примеры. 1. $y = c$ (c — постоянная).

Так как $\Delta y = c - c = 0$, то $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = 0$ и, таким образом,

$$c' = 0.$$

2. $y = \sin x$. Имеем

$$\Delta y = \sin (x + \Delta x) - \sin x = 2 \cos \left(x + \frac{\Delta x}{2}\right) \sin \frac{\Delta x}{2},$$

поэтому

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \cos \left(x + \frac{\Delta x}{2}\right) \lim_{\Delta x \to 0} \frac{\sin (\Delta x/2)}{(\Delta x/2)} = \cos x.$$

Таким образом, $(\sin x)' = \cos x$.

3. $y = \cos x$. Так как

$$\Delta y = \cos (x + \Delta x) - \cos (x) = -2 \sin \left(x + \frac{\Delta x}{2}\right) \sin \frac{\Delta x}{2},$$

то будем иметь

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = -\lim_{\Delta x \to 0} \sin \left(x + \frac{\Delta x}{2}\right) \lim_{\Delta x \to 0} \frac{\sin (\Delta x/2)}{(\Delta x/2)} = -\sin x.$$

Таким образом, $(\cos x)' = -\sin x$.

4. $y = a^x$. Имеем $\Delta y = a^{x + \Delta x} - a^x = a^x(a^{\Delta x} - 1)$, поэтому

$$\frac{\Delta y}{\Delta x} = a^x \frac{a^{\Delta x} - 1}{\Delta x},$$

откуда, в силу формулы (8.17), получаем

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = a^x \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x} = a^x \ln a.$$

Таким образом,

$$(a^x)' = a^x \ln a,$$

в частности,

$$(e^x)' = e^x.$$

Последнее равенство показывает, что число e обладает замечательным свойством: *показательная функция с основанием e имеет производную, совпадающую с самой функцией*. Этим и объясняется то обстоятельство, что в математическом анализе в качестве основания степени и основания логарифмов используется преимущественно число e. Это очень удобно, так как упрощает вычисления.

5. $y = x^n$, n — натуральное число. Используя правило возведения бинома в степень, находим

$$\Delta y = (x + \Delta x)^n - x^n = nx^{n-1}\Delta x + \frac{n(n-1)}{2}x^{n-2}\Delta x^2 + \ldots + \Delta x^n,$$

откуда при $\Delta x \neq 0$ имеем

$$\frac{\Delta y}{\Delta x} = nx^{n-1} + \frac{n(n-1)}{2}x^{n-1}\Delta x + \ldots + \Delta x^{n-1}.$$

Так как при $\Delta x \to 0$ все слагаемые правой части, содержащие множитель Δx в степени с натуральным показателем, стремятся к нулю, то $\lim\limits_{\Delta x \to 0}\frac{\Delta y}{\Delta x} = nx^{n-1}$, таким образом,

$$(x^n)' = nx^{n-1}.$$

В дальнейшем будет показано, что эта формула справедлива и тогда, когда n — произвольное действительное число.

9.2. Дифференциал функции

Определение 3. *Функция $y = f(x)$, определенная в некоторой окрестности $U(x_0)$ точки $x_0 \in \boldsymbol{R}$, называется дифференцируемой при $x = x_0$, если ее приращение в этой точке, т. е.*

$$\Delta y = f(x_0 + \Delta x) - f(x_0), \ \ x_0 + \Delta x \in U(x_0),$$

представимо в виде

$$\Delta y = A\,\Delta x + o(\Delta x), \ \ \Delta x \to 0, \tag{9.2}$$

где A — постоянная[1].

---

[1] При фиксированной точке x_0 постоянная A есть некоторое число, не зависящее от Δx; конечно, при изменении точки x_0 число A, вообще говоря, меняется.

Линейная функция $A\,\Delta x$ (от переменной Δx) называется *дифференциалом функции* f в точке x_0 и обозначается $df(x_0)$ или, короче, dy.

Таким образом,

$$dy = A\,\Delta x$$

и

$$\Delta y = dy + o(\Delta x),\ \Delta x \to 0.$$

Функция Δx определена для всех значений Δx, в частности и для $\Delta x = 0$, а функция $o(\Delta x) \underset{(9.2)}{=} \Delta y - A\,\Delta x$ определена в некоторой окрестности точки $\Delta x = 0$, в том числе и в самой этой точке, где она равна нулю:

$$o(0) \underset{(9.2)}{=} (\Delta y - A\,\Delta x)\big|_{\Delta x = 0} = 0.$$

Поэтому, согласно определению символа «o малое», существует такая функция $\varepsilon(\Delta x)$, также определенная на всей указанной окрестности точки $\Delta x = 0$ (а следовательно, в самой этой точке), что для всех $\Delta x = 0$ из этой окрестности выполняется равенство

$$o(\Delta x) = \varepsilon(\Delta x)\Delta x \qquad (9.3)$$

и

$$\lim_{\Delta x \to 0} \varepsilon(\Delta x) = 0. \qquad (9.4)$$

Предел берется по всей окрестности, поэтому отсюда следует, что

$$\varepsilon(0) = 0 \qquad (9.5)$$

(если точка, в которой берется предел функции, принадлежит множеству, по которому берется предел и этот предел существует, то функция непрерывна в этой точке, см. п. 5.5).

Заметим, что дифференциал $dy = A\,\Delta x$, как и всякая линейная функция, определен для любого значения Δx: $-\infty < \Delta x < +\infty$, в то время как приращение $\Delta y = f(x_0 + \Delta x) - f(x_0)$, естественно, можно рассматривать только для таких Δx, для которых $x_0 + \Delta x$ принадлежит области определения функции f.

Если $A \neq 0$, т. е. если $dy \not\equiv 0$, то дифференцируемость функции в точке x_0 означает, что с точностью до бесконечно малых более высокого порядка, чем приращение аргумента Δx, приращение функции Δy является линейной функцией от Δx. Используя терминологию п. 8.4, можно сказать, что главная часть приращения функции Δy в точке x_0 является линейной функцией относительно Δx; при этом приращение Δy и дифференциал dy — эквивалентные бесконечно малые при $\Delta x \to 0$ (см. п. 8.3).

Если же $A = 0$, т. е. $dy \equiv 0$, то $\Delta y = o(\Delta x)$ при $\Delta x \to 0$. Таким образом, при $A = 0$ приращение Δy является бесконечно малой более высокого порядка, чем Δx, когда $\Delta x \to 0$.

Для большей симметрии записи дифференциала приращение Δx обозначают dx и называют его *дифференциалом независимого переменного*. Таким образом, дифференциал можно записать в виде $dy = Adx$.

Пример. Найдем дифференциал функции $y = x^3$.

В этом случае

$$\Delta y = (x + \Delta x)^3 - x^3 = 3x^2\Delta x + 3x(\Delta x)^2 + (\Delta x)^3.$$

При $\Delta x \to 0$ главная линейная часть выражения, стоящего справа, равна $3x^2\Delta x$; поэтому $dy = 3x^2dx$.

Пусть $f(x_0) = y_0$. Подставив в (9.2) значения $\Delta y = f(x) - y_0$, $\Delta x = x - x_0$, $dy = A(x - x_0)$, получим

$$f(x) = y_0 + A(x - x_0) + o(x - x_0),\ x \to x_0.$$

Итак, если функция $f(x)$ дифференцируема в точке x_0, то с точностью до бесконечно малых более высокого порядка, чем $x - x_0$, вблизи x_0 она равна линейной функции; иначе говоря, в этом случае функция f в окрестности точки x_0 ведет себя «почти как линейная функция» $y_0 + A(x - x_0)$, причем погрешность при замене функции f этой линейной функцией тем меньше, чем меньше разность $x - x_0$, и, более того, отношение этой погрешности к разности $x - x_0$, т. е. относительная погрешность, стремится к нулю при $x \to x_0$.

Если функция f дифференцируема в каждой точке некоторого интервала, то ее дифференциал является функцией двух переменных — точки x и переменной dx:

$$dy = A(x)dx.$$

Выясним теперь связь между дифференцируемостью в точке и существованием производной в той же точке.

ТЕОРЕМА 1. *Для того чтобы функция f была дифференцируемой в некоторой точке x_0, необходимо и достаточно, чтобы она имела в этой точке производную; при этом*

$$dy = f'(x_0)dx.$$

Доказательство необходимости. Пусть функция f дифференцируема в точке x_0, т. е. $\Delta y = A\,\Delta x + o(\Delta x)$, $\Delta x \to 0$. Тогда

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = A + \lim_{\Delta x \to 0} \frac{o(\Delta x)}{\Delta x} = A.$$

Поэтому производная $f'(x_0)$ существует и равна A. Отсюда

$$dy = f'(x_0)dx.$$

Доказательство достаточности. Пусть существует производная $f'(x_0)$, т. е. существует предел $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(x_0)$.

Тогда

$$\frac{\Delta y}{\Delta x} = f'(x_0) + \varepsilon(\Delta x),\ \Delta x \neq 0,$$

где $\lim\limits_{\substack{\Delta x \to 0 \\ \Delta x \neq 0}} \varepsilon(\Delta x) = 0$ и, следовательно, для $\Delta x \neq 0$ справедливо равенство

$$\Delta y = f'(x_0)\,\Delta x + \varepsilon(\Delta x)\,\Delta x.$$

Полагая здесь $\varepsilon(0) = 0$, получаем, что в некоторой окрестности точки x_0 имеет место равенство

$$\Delta y = f'(x_0)\,\Delta x + o(\Delta x),\quad \Delta x \to 0,$$

т. е. равенство (9.2) при $A = f'(x_0)$. Таким образом, функция f дифференцируема в точке x_0. □

Подчеркнем, что в теореме 1 речь идет о конечной производной.

Таким образом, *дифференцируемость* функции $f(x)$ в точке x_0 *равносильна существованию* в этой точке *конечной производной* $f'(x_0)$.

Из доказанного следует, что коэффициент A в определении дифференциала (см. (9.2)) определен однозначно, а именно $A = f'(x_0)$; тем самым и дифференциал функции в данной точке определен однозначно. Это, впрочем, вытекает также из леммы п. 8.4 о единственности главной части вида $A(x - x_0)^k$ бесконечно малой функции при $x \to x_0$.

В силу теоремы 1, $y' = \frac{dy}{dx}$. Правая часть представляет собой дробь, числитель которой — дифференциал функции, а знаменатель — дифференциал аргумента.

Формулу (9.2) для приращения функции, согласно теореме 1 и формуле (9.3), можно записать в виде

$$\Delta y = f'(x_0)\,\Delta x + \varepsilon(\Delta x)\,\Delta x, \tag{9.6}$$

где выполняется условие (9.4), следовательно, и условие (9.5):

$$\lim_{\Delta x \to 0} \varepsilon(\Delta x) = 0,\ \ \varepsilon(0) = 0, \tag{9.7}$$

т. е. $\varepsilon(\Delta x)$ — непрерывная в нуле функция.

Формула $dy = f'(x_0)dx$ позволяет находить дифференциалы функций, если известны их производные. Так, например, используя производные, найденные в п. 9.1, получаем:

$$dc = 0\ (c \text{ — постоянная}),\quad d\cos x = -\sin x\,dx,$$

$$d\sin x = \cos x\,dx,\quad da^x = a^x \ln a\,dx,$$

в частности $de^x = e^x\,dx$,

$$dx^n = nx^{n-1}\,dx\ (n \text{ — натуральное число}).$$

В заключение выясним связь между дифференцируемостью и непрерывностью в данной точке.

ТЕОРЕМА 2. *Если функция f дифференцируема в некоторой точке, то она и непрерывна в этой точке.*

СЛЕДСТВИЕ. *Если функция в некоторой точке имеет производную, то она непрерывна в этой точке.*

Доказательство. Пусть функция f дифференцируема в точке x_0, т. е. в этой точке $\Delta y = A\,\Delta x + o(\Delta x)$ при $\Delta x \to 0$.

Тогда

$$\lim_{\Delta x \to 0} \Delta y = A \lim_{\Delta x \to 0} \Delta x + \lim_{\Delta x \to 0} o(\Delta x) = 0,$$

что и означает непрерывность функции f при $x = x_0$. □

Следствие непосредственно вытекает из теорем 1 и 2. Обратим внимание на то, что если функция имеет в точке бесконечную производную, то она может быть разрывной в этой точке.

УПРАЖНЕНИЕ 1. Построить пример функции, имеющей в некоторой точке бесконечную производную и разрывную в этой точке.

Заметим, что утверждение, обратное теореме 2, неверно, т. е. из непрерывности функции f в данной точке не следует ее дифференцируемость или, что равносильно (см. теорему 1), существование производной в этой точке.

Приведем примеры, подтверждающие это.

Примеры. 1. Функция $f(x) = |x|$, очевидно, непрерывна в точке $x = 0$ (как и во всех других), но не имеет в этой точке производной.

В самом деле, при $x \geqslant 0$ имеем $y = |x| = x$, поэтому для точки $x_0 = 0$ получим $\Delta y = \Delta x$. Следовательно,

$$f'_+(0) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = 1.$$

Аналогично, при $x \leqslant 0$ имеем $y = |x| = -x$, поэтому для точки $x_0 = 0$ в этом случае получим $\Delta y = -\Delta x$. Следовательно,

$$f'_-(0) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = -1.$$

Тем самым доказано, что функция $f(x) = |x|$ не имеет при $x = 0$ производной, однако в этой точке существуют как левая, так и правая производные.

Отметим еще, что при $x > 0$ имеет место равенство $(|x|)' = x' = 1$, а при $x < 0$ соответственно $(|x|)' = (-x') = -1$, поэтому для любого $x \neq 0$ справедлива формула

$$|x|' = \operatorname{sign} x.$$

Следующий пример показывает, что у функции в точке непрерывности может не быть никакой односторонней производной.

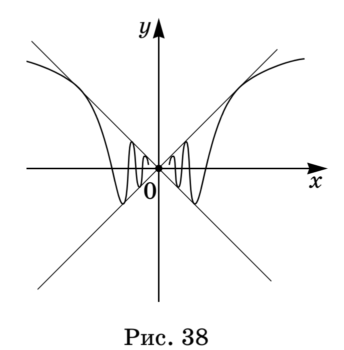


Рис. 38

2. Пусть (рис. 38)

$$f(x) = \begin{cases} x \sin \frac{1}{x} & \text{при } x \neq 0, \\ 0 & \text{при } x = 0. \end{cases}$$

Тогда в точке $x = 0$ имеем $\Delta y = \Delta x \sin \frac{1}{\Delta x}$, откуда $|\Delta y| \leqslant |\Delta x|$, и поэтому $\lim\limits_{\Delta x \to 0} \Delta y = 0$, т. е. рассматриваемая функция непрерывна при $x = 0$. Вместе с тем $\frac{\Delta y}{\Delta x} = \sin \frac{1}{\Delta x}$, и так как функция $\sin \frac{1}{x}$ не имеет в точке $x = 0$ предела ни слева, ни справа (см. пример 2 в п. 5.4), то у функции $f(x)$ не существует односторонних производных при $x = 0$.

УПРАЖНЕНИЯ. 2. Ввести понятие дифференцируемости функции справа (слева) в данной точке и доказать, что дифференцируемость справа (слева) в данной точке эквивалентна существованию в этой точке конечной производной справа (слева).

3. Доказать, что если функция имеет в некоторой точке левостороннюю (правостороннюю) производную, то в этой точке функция непрерывна слева (справа).

Если функция f имеет производную в каждой точке некоторого промежутка (дифференцируема в каждой точке этого промежутка), то говорят, что функция f *имеет производную* или что она *дифференцируема на* указанном *промежутке*.

9.3. Геометрический смысл производной и дифференциала

Понятия производной и дифференциала функции в данной точке связаны с понятием касательной к графику функции в этой точке. Чтобы выяснить эту связь, определим прежде всего касательную.

Пусть функция $y = f(x)$ определена на интервале (a, b) и непрерывна в точке $x_0 \in (a, b)$. Пусть $y_0 = f(x_0)$, $M_0 = (x_0, y_0)$, $x_0 + \Delta x \in (a, b)$, $\Delta y = f(x_0 + \Delta x) - f(x_0)$,

$$M = (x_0 + \Delta x, y_0 + \Delta y).$$

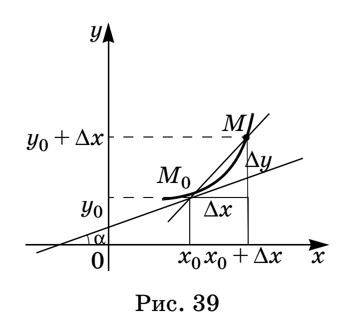


Рис. 39

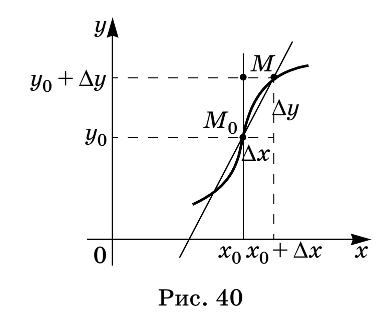


Рис. 40

Проведем секущую M_0M (рис. 39). Она имеет уравнение

$$y = k(\Delta x)(x - x_0) + y_0, \tag{9.8}$$

где

$$k(\Delta x) = \frac{\Delta y}{\Delta x}. \tag{9.9}$$

Покажем, что при $\Delta x \to 0$ расстояние $|M_0M|$ от точки M_0 до точки M стремится к нулю (в этом случае говорят, что точка M стремится к точке M_0, и пишут $M \to M_0$). Действительно, в силу непрерывности функции f, при $x = x_0$ имеем $\lim\limits_{\Delta x \to 0} \Delta y = 0$. Следовательно, при $\Delta x \to 0$

$$|M_0M| = \sqrt{\Delta x^2 + \Delta y^2} \to 0.$$

Определение 4. *Если существует конечный предел* $\lim\limits_{\Delta x \to 0} k(\Delta x) = k_0$, *то прямая, уравнение которой*

$$y = k_0(x - x_0) + y_0 \tag{9.10}$$

получается из уравнения $y = k(\Delta x)(x - x_0) + y_0$ *при* $\Delta x \to 0$ (рис. 39), *называется* (*наклонной*) *касательной к графику функции* f *в точке* (x_0, y_0).

Если $\lim\limits_{\Delta x \to 0} k(\Delta x) = \infty$, *то прямая* (рис. 40), *уравнение которой*

$$x = x_0 \tag{9.11}$$

получается при $\Delta x \to 0$ *из уравнения секущей, записанного в виде* $\frac{y}{k(\Delta x)} = x - x_0 + \frac{y}{k(\Delta x)}$, *называется* (*вертикальной*) *касательной к графику функции* f *в точке* (x_0, y_0).

Прямые (9.10) в случае конечного предела $\lim\limits_{\Delta x \to 0} k(\Delta x)$ и (9.11) в случае, когда этот предел бесконечен, называются *предельными положениями прямой* (9.8). В силу этого, данное выше определение касательной к графику функции можно перефразировать следующим образом.

Предельное положение секущей M_0M *при* $\Delta x \to 0$, *или, что то же, при* $M \to M_0$, *называется касательной к графику функции* f *в точке* M_0.

Заметим теперь, что, в силу равенства (9.9), существование конечного предела $\lim\limits_{\Delta x \to 0} k(\Delta x) = \lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$ означает существование конечной производной $f'(x_0) = k$. Следовательно, если у функции f в точке x_0 существует конечная производная, то уравнение касательной к графику функции в точке $(x_0, f(x_0))$ имеет вид

$$y = f'(x_0)(x - x_0) + y_0, \qquad (9.12)$$

где $y_0 = f(x_0)$. Если же $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \infty$, т. е. $f'(x_0) = \infty$, то, в силу (9.9), $\lim\limits_{\Delta x \to 0} k(\Delta x) = \infty$ и, следовательно (см. (9.11)), уравнение касательной имеет вид $x = x_0$.

Как известно из аналитической геометрии, коэффициент $f'(x_0)$ в уравнении (9.12) равен тангенсу угла (см. рис. 39), который рассматриваемая прямая образует с положительным направлением оси Ox: $f'(x_0) = \operatorname{tg} \alpha$, т. е. производная функции в некоторой точке равна тангенсу угла между касательной в соответствующей точке графика функции и осью абсцисс.

Первое слагаемое в правой части уравнения (9.12), т. е. выражение $f'(x_0)(x - x_0) = f'(x_0)\,\Delta x$, $\Delta x = x - x_0$, является дифференциалом dy функции f в точке x_0. Следовательно, в силу равенства (9.12),

$$y - y_0 = dy,$$

где y — текущая ордината касательной. Таким образом, дифференциал функции в данной точке равен приращению ординаты касательной в соответствующей точке графика функции.

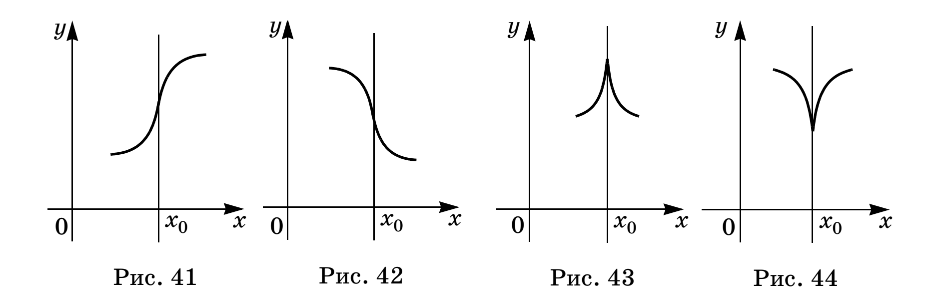


Рис. 41 Рис. 42 Рис. 43 Рис. 44

З а м е ч а н и е. Если в точке x_0 существует бесконечный предел $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \infty$, то может оказаться, что он будет равным $+\infty$ или $-\infty$. В этом случае при $x = x_0$ существует бесконечная производная $y' = +\infty$ или $y' = -\infty$ и график функции $y = f(x)$ в окрестности точки x_0 имеет вид, схематически изображенный на рисунках 41 и 42.

Возможен также и случай, когда предел $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \infty$ не является бесконечностью определенного знака и, следовательно, в этой точке не существует ни конечной, ни бесконечной определенного знака производной, а лишь $f'(x_0) = \infty$. Это может, например, случиться, если в точке x_0 существуют односторонние бесконечные производные разного знака. Тогда в окрестности точки x_0 график функции имеет вид, схематически показанный на рисунках 43 и 44.

Пример. Найдем касательную к параболе $y = x^2$ в точке с координатами (1, 1).

Согласно п. 9.1 (см. пример 5), $y' = 2x$, поэтому $y'|_{x=1} = 2$. В силу формулы (9.12), искомая касательная имеет уравнение $y = 2(x - 1) + 1$, т. е. $y = 2x - 1$.

Если функция f дифференцируема в точке x_0, то, подставляя в формулу (9.5) $A = f'(x_0)$ (см. теорему 1 настоящего параграфа), имеем

$$f(x) = y_0 + f'(x_0)(x - x_0) + o(x - x_0),\ x \to x_0$$

и, значит, согласно (9.12) ($y_{\text{кас}} = f'(x_0)(x - x_0) + y_0$), получим

$$f(x) - y_{\text{кас}} = o(x - x_0),\ x \to x_0.$$

Таким образом, наклонная касательная к графику функции обладает тем свойством, что разность ординат графика и этой касательной есть величина бесконечно малая более высокого порядка при $x \to x_0$ по сравнению с приращением аргумента.

Обратно: если существует невертикальная прямая

$$y_{\text{пр}} = A(x - x_0) + y_0, \tag{9.13}$$

проходящая через точку (x_0, y_0) и такая, что

$$f(x) - y_{\text{пр}} = o(x - x_0),\ x \to x_0, \tag{9.14}$$

то эта прямая является касательной к графику функции в точке (x_0, y_0). Действительно, в этом случае

$$f(x) - [A(x - x_0) + y_0] = o(x - x_0),$$

т. е.

$$\Delta y = f(x) - y_0 = A(x - x_0) + o(x - x_0),\ x \to x_0;$$

следовательно, функция f дифференцируема в точке x_0 (см. (9.2)) и $A = f'(x_0)$ (см. теорему 1), т. е. указанная прямая совпадает с касательной (9.12).

Таким образом, *условие* (9.14) *необходимо и достаточно для того*, *чтобы прямая* (9.13) *являлась наклонной касательной к графику функции* $f(x)$ *в точке* (x_0, y_0). Отсюда, в частности, следует, что она единственна (последнее вытекает, например, из того, что дифференциал функции единствен, или из того, что касательная к графику функции в данной точке единственна).

9.4. Физический смысл производной и дифференциала

Пусть функция $f(x)$ определена в некоторой окрестности точки x_0. Воспользуемся, как и выше, обозначениями $\Delta x = x - x_0$, $\Delta y = f(x_0 + \Delta x_0) - f(x_0)$. Пусть для определенности $\Delta x > 0$. Отношение $\frac{\Delta y}{\Delta x}$, равное изменению переменной y на отрезке $[x_0, x_0 + \Delta x]$, отнесенному к единице измерения переменной x, естественно назвать *значением средней скорости изменения* y на отрезке $[x_0, x_0 + \Delta x]$ относительно x. При стремлении Δx к нулю, т. е. при стягивании отрезка

$[x_0, x_0 + \Delta x]$ к точке x_0, отношение $\frac{\Delta y}{\Delta x}$ определяет значение средней скорости изменения переменной y относительно переменной x во все меньшем и меньшем отрезке, содержащем точку x_0. Все сказанное, конечно, справедливо и при $\Delta x < 0$ для отрезка $[x_0 + \Delta x, x_0]$.

Предел $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$, если он существует, т. е. производную $f'(x_0)$, естественно поэтому назвать *скоростью* изменения переменной y относительно переменной x в точке x_0.

Заметим, что если в точке x_0 существует производная $f'(x_0)$, то, рассматривая предел средних скоростей изменения y относительно x на отрезках $[x_0 - \Delta x, x_0 + \Delta x]$ $(\Delta x > 0)$, содержащих точку x_0 внутри себя в качестве центра, при стягивании их к точке x_0 (при $\Delta x \to 0$) мы получим в пределе то же значение скорости изменения y относительно x в точке x_0, и т. е. $f'(x_0)$. Действительно, значение средней скорости изменения переменной y относительно x на отрезке $[x_0 - \Delta x, x_0 + \Delta x]$ равно $\frac{f(x_0 + \Delta x) - f(x_0 - \Delta x)}{2\Delta x}$ (частному от деления изменения функции на длину отрезка, на котором произошло это изменение); отсюда

$$\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0 - \Delta x)}{2\Delta x} =$$

$$= \frac{1}{2}\left[\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} + \lim_{\Delta x \to 0} \frac{f(x_0 - \Delta x) - f(x_0)}{-\Delta x}\right] = f'(x_0).$$

Интересно заметить, что разностное отношение $\frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x}$ в известном смысле лучше приближает значение производной f' в точке x, чем $\frac{f(x + \Delta x) - f(x)}{\Delta x}$ (см. п. 62.6).

На интерпретации производной как скорости изменения одной величины относительной другой и основано применение производной к изучению физических явлений.

Применение же дифференциала основано на том, что замена приращения функции ее дифференциалом позволяет заменить любую дифференцируемую в точке x_0 функцию линейной функцией в достаточно малой окрестности точки x_0, т. е. считать, что процесс изменения зависимости пере-

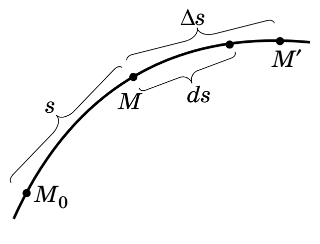


Рис. 45

менной «в малом» происходит линейно относительно аргумента. Иначе говоря, можно считать, что изменение функции прямо пропорционально изменению аргумента или, как говорят, упомянутый процесс «в малом» происходит равномерно. Получающаяся при такой замене погрешность оказывается бесконечно малой более высокого порядка, чем приращение аргумента.

Примеры. 1. Пусть $s = s(t)$ — закон движения материальной точки[1] (рис. 45); s — длина пути, отсчитываемая вдоль траектории от некоторой начальной точки M_0; t — время. Пусть M — положение точки в момент времени t, а M' — в момент $t + \Delta t$ и Δs — длина пути от M до M', т. е. $\Delta s = s(t + \Delta t) - s(t)$.

Отношение $\frac{\Delta s}{\Delta t}$ называется в механике числовым значением средней скорости движения на участке от M до M', а $\lim\limits_{\Delta x \to 0} \frac{\Delta s}{\Delta t} = v$ — числовым значением скорости в точке M или числовым значением *мгновенной скорости* в момент времени t; таким образом, $v = \frac{ds}{dt}$.

По определению дифференциала, $ds = vdt$; следовательно, дифференциал пути равен расстоянию, которое прошла бы точка за промежуток времени от момента t до $t + \Delta t$, если бы она двигалась равномерно со скоростью, равной мгновенной скорости (в смысле ее числового значения) точки в момент t. Величина же Δs действительного перемещения точки равна $\Delta s = ds + o(\Delta t)$.

Мы видим таким образом, что с точки зрения механики замена Δs на ds означает, что мы считаем движение на рассматриваемом участке равномерным (в смысле числового значения скорости[2]).

---

[1] Не следует путать закон движения точки с уравнением ее траектории, которое имеет вид $\boldsymbol{r} = \boldsymbol{r}(t)$, где $\boldsymbol{r}$ — радиус-вектор движущейся точки.

[2] Следует иметь в виду, что скорость — вектор и поэтому характеризуется не только значением, но и направлением.

2. Пусть $q = q(t)$ — количество электричества, протекающее через поперечное сечение проводника; t — время; Δt — некоторый промежуток времени; $\Delta q = q(t + \Delta t) - q(t)$ — количество электричества, протекающее через указанное сечение за промежуток времени от момента t до момента $t + \Delta t$. Тогда $\frac{\Delta q}{\Delta t}$ называется *средней силой тока* за промежуток времени Δt и обозначается $I_{\text{ср}}$, а предел $\lim\limits_{\Delta t \to 0} I_{\text{ср}} = \lim\limits_{\Delta t \to 0} \frac{\Delta q}{\Delta t}$ — *силой тока в данный момент времени* t или *мгновенной силой тока* и обозначается I. Таким образом, $I = \frac{dq}{dt}$. Дифференциал $dq = I\Delta t$ равен количеству электричества, протекшего через поперечное сечение проводника за промежуток времени Δt, если сила тока была бы постоянной и равной силе тока в момент t. Как всегда, $\Delta q - dq = o(\Delta t)$, $\Delta t \to 0$.

$m(x)$ Δm

3 x $x + \Delta x$ l x

Рис. 46

3. Пусть дан неоднородный стержень[1] длины l и пусть $m = m(x)$ — масса стержня длины x, $0 \leqslant x \leqslant l$, отмеряемой от одного фиксированного конца (рис. 46). Тогда $\Delta m = m(x + \Delta x) - m(x)$ — масса части стержня, ограниченной точками, расположенными соответственно на расстоянии x и $x + \Delta x$ от указанного конца. Величина $\frac{\Delta m}{\Delta x}$ называется *средней линейной плотностью* стержня на указанном участке и обозначается $\rho_{\text{ср}}$. Предел $\lim\limits_{\Delta x \to 0} \rho_{\text{ср}} = \lim\limits_{\Delta x \to 0} \frac{\Delta m}{\Delta x}$ называется *линейной плотностью* стержня в данной точке и обозначается ρ. Таким образом, $\rho = \frac{dm}{dx}$.

Если плотность ρ постоянна, то стержень является однородным.

Для произвольного, вообще говоря, неоднородного стержня дифференциал $dm = \rho\, \Delta x$ равен массе однородного стержня длины Δx с постоянной плотностью ρ, равной плотности рассматриваемого стержня в данной точке.

[1] Стержень называется однородным, если два любых его участка одинаковой длины имеют одинаковую массу, и неоднородным — в противном случае.

Как показывает этот пример, интерпретируя производную как скорость, мы должны понимать это в широком смысле слова. Например, плотность стержня тоже «скорость», а именно скорость изменения массы с изменением длины.

9.5. Правила вычисления производных, связанные с арифметическими действиями над функциями

Получим теперь формулы для производных суммы, произведения и частного функций.

ТЕОРЕМА 3. *Пусть функция $y_1 = f_1(x)$ и $y_2 = f_2(x)$ определены в окрестности точки $x_0 \in \boldsymbol{R}$ и имеют в самой точке x_0 производные, тогда и их сумма $f_1(x) + f_2(x)$, произведение $f_1(x)f_2(x)$, а если $f_2(x) \neq 0$, то и частное $\frac{f_1(x)}{f_2(x)}$ имеют в точке x_0 производные, причем*

$$(y_1 + y_2)' = y_1' + y_2', \tag{9.15}$$

$$(y_1 y_2)' = y_1' y_2 + y_1 y_2', \tag{9.16}$$

$$\left(\frac{y_1}{y_2}\right)' = \frac{y_1' y_2 - y_1 y_2'}{y_2^2} \tag{9.17}$$

(в формулах (9.15), (9.16) и (9.17) $x = x_0$).

СЛЕДСТВИЕ 1. *Если функция $y = f(x)$ имеет производную в точке x_0 и $c \in \boldsymbol{R}$, то функция $cf(x)$ также имеет в этой точке производную, причем*

$$(cy)' = cy' \quad (x = x_0). \tag{9.18}$$

СЛЕДСТВИЕ 2. *Если функции $y_k = f_k(x)$, $k = 1, 2, \ldots, n$, имеют в точке x_0 производные, то всякая их линейная комбинация также имеет в этой точке производную, причем*

$$(c_1 y_1 + \ldots + c_n y_n)' = c_1 y_1' + \ldots + c_n y_n', \quad c_k \in \boldsymbol{R},\ k = 1, 2, \ldots, n.$$

Доказательство теоремы. Пусть функции $y_1 = f_1(x)$ и $y_2 = f_2(x)$ определены в окрестности $U(x_0)$ точки x_0, $x_0 + \Delta x \in U(x_0)$ и $\Delta y_1 = f_1(x_0 + \Delta x) - f_1(x_0)$, $\Delta y_2 = f_2(x_0 + \Delta x) - f_2(x_0)$. Для простоты записи будем иногда опускать обозначение аргумента, рассматривая при этом приращения функций только в точке x_0.

Если $y = y_1 + y_2$, то

$$\Delta y = (y_1 + \Delta y_1 + y_2 + \Delta y_2) - (y_1 + y_2) = \Delta y_1 + \Delta y_2,$$

откуда при $\Delta x \neq 0$ получим

$$\frac{\Delta y}{\Delta x} = \frac{\Delta y_1}{\Delta x} + \frac{\Delta y_2}{\Delta x}.$$

Переходя здесь к пределу при $\Delta x \to 0$ и замечая, что в силу существования производных функций y_1 и y_2 в точке x_0 предел правой части этого равенства существует и равен $y_1' + y_2'$, получим, что существует и предел его левой части, т. е. существует производная y', причем $y' = y_1' + y_2'$, т. е. формула (9.15) доказана.

Если $y = y_1 y_2$, то аналогичным образом будем последовательно иметь

$$\Delta y = (y_1 + \Delta y_1)(y_2 + \Delta y_2) - y_1 y_2 =$$

$$= \Delta y_1 \cdot y_2 + y_1 \cdot \Delta y_2 + \Delta y_1 \cdot \Delta y_2,$$

$$\frac{\Delta y}{\Delta x} = \frac{\Delta y_1}{\Delta x} y_2 + y_1 \frac{\Delta y_2}{\Delta x} + \frac{\Delta y_1}{\Delta x} \Delta y_2.$$

Из существования производной $f_2'(x_0)$ следует непрерывность функции f_2 в точке x_0: $\lim\limits_{\Delta x \to 0} \Delta y_2 = 0$; кроме того, $\lim\limits_{\Delta x \to 0} \frac{\Delta y_1}{\Delta x} = y_1'$, $\lim\limits_{\Delta x \to 0} \frac{\Delta y_2}{\Delta x} = y_2'$. Поэтому, перейдя к пределу при $\Delta x \to 0$, из полученного равенства имеем

$$y' = y_1' y_2 + y_1 y_2',$$

т. е. формула (9.16) доказана.

Наконец, если $y = \frac{y_1}{y_2}$ и $f_2(x_0) \neq 0$, то

$$\Delta y = \frac{y_1 + \Delta y_1}{y_2 + \Delta y_2} - \frac{y_1}{y_2} = \frac{\Delta y_1 \cdot y_2 - y_1 \cdot \Delta y_2}{(y_2 + \Delta y_2) y_2}, \quad \frac{\Delta y}{\Delta x} = \frac{\frac{\Delta y_1}{\Delta x} y_2 - y_1 \frac{\Delta y_2}{\Delta x}}{(y_2 + \Delta y_2) y_2}.$$

Отсюда при $\Delta x \to 0$, вспомнив снова, что из существования производной следует непрерывность функции, и, следова-

тельно, $\lim\limits_{\Delta x \to 0} \Delta y_2 = 0$, получим $y' = \dfrac{y'_1\, y_2 - y_1 y'_2}{y_2^2}$, т. е. формула (9.17) также доказана.

Следствие 1 сразу вытекает из (9.16), если вспомнить, что $c' = 0$ (см. пример 1 в п. 9.1), а следствие 2 сразу получается из формул (9.15) и (9.18) методом математической индукции. □

З а м е ч а н и е. Используя свойства бесконечных пределов, относящиеся к арифметическим действиям над функциями (см. п. 5.10), можно установить и соответствующие свойства бесконечных производных. Например, если существует конечная производная $y'_1(x_0)$ и бесконечная (определенного знака) производная $y'_2(x_0)$, то у функции $y(x) \overset{\text{def}}{=} y_1(x) + y_2(x)$ в точке x_0 существует бесконечная производная того же знака. Например, если $y'_2(x_0) = +\infty$, то $y'(x_0) = +\infty$. Действительно, $\Delta y = \Delta y_1 + \Delta y_2$. Поэтому, если существует конечный предел $\lim\limits_{\Delta x \to 0} \dfrac{\Delta y_1}{\Delta x}$, а $\lim\limits_{\Delta x \to 0} \dfrac{\Delta y^2}{\Delta x} =$ $= +\infty$, то

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \left(\frac{\Delta y_1}{\Delta x} + \frac{\Delta y_2}{\Delta x}\right) = \lim_{\Delta x \to 0} \frac{\Delta y_1}{\Delta x} + \lim_{\Delta x \to 0} \frac{\Delta y_2}{\Delta x} = +\infty,$$

т. е. $y'(x_0) = +\infty$.

Примеры. 1. Пусть $y = e^x \sin x - 2x^2 \cos x$; в силу формул (9.15), (9.16) и (9.18), имеем

$$y' = (e^x \sin x)' - 2(x^2 \cos x)' =$$
$$= e^x \sin x + e^x \cos x - 2(2x \cos x - x^2 \sin x).$$

2. Пусть $y = \operatorname{tg} x$; так как $\operatorname{tg} x = \dfrac{\sin x}{\cos x}$, то по формуле (9.17) получаем

$$y' = \left(\frac{\sin x}{\cos x}\right)' = \frac{\cos x \cos x - \sin x\,(-\sin x)}{\cos^2 x} = \frac{1}{\cos^2 x}.$$

Таким образом,

$$(\operatorname{tg} x)' = \frac{1}{\cos^2 x}.$$

3. Аналогично, для $y = \operatorname{ctg} x$

$$y' = \left(\frac{\cos x}{\sin x}\right)' = \frac{(-\sin x)\sin x - \cos x \cos x}{\sin^2 x} = -\frac{1}{\sin^2 x},$$

т. е.

$$(\operatorname{ctg} x)' = -\frac{1}{\sin^2 x}.$$

Свойства (9.15)—(9.18) переносятся и на дифференциалы функций. При тех же предположениях относительно дифференцируемости в точке x_0 имеем:

$$d(y_1 + y_2) = dy_1 + dy_2;\ \ d(y_1 y_2) = y_2 dy_1 + y_1 dy_2;$$
$$d(cy) = c dy;\ \ d\left(\frac{y_1}{y_2}\right) = \frac{y_2 dy_1 - y_1 dy_2}{y_2^2}.$$

Вычислим, например, дифференциал произведения $y = y_1 y_2$:

$$dy = y'dx = (y_1 y_2)'dx = y_1' y_2 dx + y_1 y_2' dx = y_2 dy_1 + y_1 dy_2,$$

так как $y_1' dx = dy_1$, $y_2' dx = dy_2$.

Аналогично доказываются и остальные формулы.

9.6. Производная обратной функции

ТЕОРЕМА 4. *Пусть функция $y = f(x)$ непрерывна и строго монотонна в некоторой окрестности точки x_0 и пусть при $x = x_0$ существует производная $\frac{df(x_0)}{dx} \neq 0$; тогда обратная функция $x = f^{-1}(y)$ имеет производную в точке $y_0 = f(x_0)$, причем*

$$\frac{df^{-1}(y_0)}{dy} = \frac{1}{\frac{df(x_0)}{dx}}, \tag{9.20}$$

т. е. производная обратной функции равна обратной величине производной данной функции.

Доказательство. Зафиксируем какую-то окрестность точки x_0, на которой функция f определена, непрерывна и строго монотонна, и будем рассматривать f только в этой окрестности. Тогда, как доказано ранее (см. п. 6.3), обратная функция определена и непрерывна на некотором интервале, содержащем

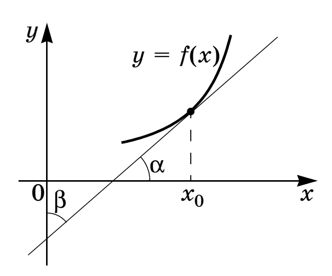


Рис. 47

точку y_0 и являющемся образом указанной выше окрестности точки x_0. Поэтому если $\Delta x = x - x_0$, $\Delta y = y - y_0$, $y = f(x)$, то $\Delta x \to 0$ равносильно $\Delta y \to 0$ в том смысле, что $\lim\limits_{\Delta x \to 0} \Delta y = 0$ (для функции f) и $\lim\limits_{\Delta y \to 0} \Delta x = 0$ (для функции f^{-1}).

Для любых $\Delta x \neq 0$, $\Delta y \neq 0$ имеем $\dfrac{\Delta x}{\Delta y} = \dfrac{1}{\Delta y / \Delta x}$. При $\Delta x \to 0$ (или, что то же, в силу сказанного выше, при $\Delta y \to 0$) предел правой части существует, значит, существует и предел левой части, причем

$$\lim_{\Delta y \to 0} \frac{\Delta x}{\Delta y} = \lim_{\Delta x \to 0} \frac{\Delta x}{\Delta y} = \frac{1}{\lim\limits_{\Delta x \to 0} \dfrac{\Delta y}{\Delta x}} = \frac{1}{\dfrac{df(x_0)}{dx}}.$$

Но $\lim\limits_{\Delta y \to 0} \dfrac{\Delta x}{\Delta y} = \dfrac{df^{-1}(y_0)}{dy}$, поэтому $\dfrac{df^{-1}(y_0)}{dy} = \dfrac{1}{\dfrac{df(x_0)}{dx}}$. □

Этой теореме можно дать наглядную геометрическую интерпретацию (рис. 47). Как известно, $\dfrac{df(x_0)}{dx} = \operatorname{tg} \alpha$, где α — значение угла, образуемого касательной графика функции f в точке (x_0, y_0) с положительным направлением оси Ox, а $\dfrac{df^{-1}(y_0)}{dy} = \operatorname{tg} \beta$, где β — значение угла, образованного той же касательной с осью Oy.

Очевидно, $\beta = \pi/2 - \alpha$, поэтому

$$\frac{df^{-1}(y_0)}{dy} = \operatorname{tg} \beta = \frac{1}{\operatorname{ctg} \beta} = \frac{1}{\operatorname{ctg}\left(\dfrac{\pi}{2} - \alpha\right)} = \frac{1}{\operatorname{tg} \alpha} = \frac{1}{\dfrac{df(x_0)}{dx}}.$$

УПРАЖНЕНИЯ. 4. Доказать, что если функция $y = f(x)$ непрерывна и строго монотонна в некоторой окрестности точки x_0 и если в этой точке существует производная и $\dfrac{df(x_0)}{dx} = 0$, то обратная функция $f^{-1}(y)$ имеет в точке $y_0 = f(x_0)$ бесконечную производную; следовательно, если считать условно, что $\dfrac{1}{0} = \infty$, то формула (9.20) справедлива и в этом случае.

5. Сформулируйте и докажите аналог теоремы 3 для односторонних производных (конечных и бесконечных).

Примеры. 1. $y = \arcsin x,\ \ x = \sin y,\ \ -\frac{\pi}{2} \leqslant y \leqslant \frac{\pi}{2},\ \ -1 \leqslant$ $\leqslant x \leqslant 1$. Применяя формулу (9.20), получаем

$$\frac{dy}{dx} = (\arcsin x)' = \frac{1}{\frac{dx}{dy}} = \frac{1}{\cos y}.$$

Так как $-\frac{\pi}{2} \leqslant y \leqslant \frac{\pi}{2}$, то $\cos y > 0$, поэтому $\cos y = \sqrt{1 - \sin^2 y} =$ $= \sqrt{1 - x^2}$. Таким образом, $(\arcsin x)' = \frac{1}{\sqrt{1 - x^2}}$.

2. $y = \arccos x,\ \ x = \cos y,\ \ 0 \leqslant y \leqslant \pi,\ \ -1 \leqslant x \leqslant 1$. Аналогично предыдущему примеру имеем:

$$\frac{dy}{dx} = (\arccos x)' = \frac{1}{\frac{dx}{dy}} = -\frac{1}{\sin y} = -\frac{1}{\sqrt{1 - \cos^2 y}} = -\frac{1}{\sqrt{1 - x^2}},$$

т. е. $(\arccos x)' = \frac{1}{\sqrt{1 - x^2}}$.

3. $y = \operatorname{arctg} x$, $x = \operatorname{tg} y$, $-\frac{\pi}{2} < y < \frac{\pi}{2}$, $-\infty < x < +\infty$. Имеем:

$$\frac{dy}{dx} = (\operatorname{arctg} x)' = \frac{1}{\frac{dy}{dx}} = \cos^2 y = \frac{1}{1 + \operatorname{tg}^2 y} = \frac{1}{1 + x^2};$$

итак, $(\operatorname{arctg} x)' = \frac{1}{1 + x^2}$.

4. $y = \operatorname{arctg} x$, $x = \operatorname{ctg} y$, $0 < y < \pi$, $-\infty < x < \infty$. В этом случае

$$\frac{dy}{dx} = (\operatorname{arcctg} x)' = \frac{1}{\frac{dx}{dy}} = -\sin^2 y = -\frac{1}{1 + \operatorname{ctg}^2 y} = -\frac{1}{1 + x^2},$$

т. е. $(\operatorname{arcctg} x)' = -\frac{1}{1 + x^2}$.

5. Если $y = \log_a x$, $x = a^y$, $a > 0$, $a \neq 1$, $x > 0$, $-\infty < y < +\infty$, то

$$\frac{dy}{dx} = (\log_a x)' = \frac{1}{\frac{dx}{dy}} = \frac{1}{a^y \ln a} = \frac{1}{x \ln a},$$

т. е.

$$(\log_a x)' = \frac{1}{x \ln a};$$

в частности, при $a = e$ имеем $(\ln x)' = \frac{1}{x}$.

9.7. Производная и дифференциал сложной функции

ТЕОРЕМА 5. *Пусть функция $y = f(x)$ имеет производную в точке x_0, а функция $z = F(y)$ имеет производную в точке $y_0 = f(x_0)$. Тогда сложная функция $\Phi(x) = F[f(x)]$ также имеет производную при $x = x_0$, причем*

$$\Phi'(x_0) = F'(y_0)f'(x_0). \tag{9.21}$$

Если сложную функцию Φ обозначить символом $\Phi = F \circ f$ (см. п. 5.2), то формулу (9.21) можно записать в виде

$$(F \circ f)'(x_0) = F'(f(x_0))f'(x_0).$$

Следует обратить внимание на то, что утверждение о существовании в точке x_0 производной сложной функции $F[f(x)]$ содержит предположение о том, что рассматриваемая сложная функция имеет смысл, т. е. определена в некоторой окрестности точки x_0.

Опуская значение аргумента и используя запись производной с помощью дифференциалов, равенство (9.21) можно переписать в виде

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}.$$

Доказательство. Прежде всего, в силу самого определения производной, функция F определена в некоторой окрестности $V(y_0)$ точки y_0, а так как из существования производной $f'(x_0)$ следует непрерывность функции f, то для указанной окрестности $V(y_0)$ существует такая окрестность $U(x_0)$ точки x_0, что $f(U(x_0)) \subset V(y_0)$, и, следовательно, для всех $x \in U(x_0)$ имеет смысл сложная функция $F(f(x))$.

Положим, как всегда, $\Delta x = x - x_0$, $\Delta y = y - y_0$, $\Delta z = F(y) - F(y_0)$. Функция F имеет в точке y_0 производную и поэтому дифференцируема в этой точке (см. п. 9.2). Это означает, что ее приращение Δz при всех Δy, принадлежащих некоторой окрестности точки $\Delta y = 0$ (в том числе и при $\Delta y = 0$), представимо (см. формулы (9.6) и (9.7)) в виде

$$\Delta z = F'(y_0)\,\Delta y + \varepsilon(\Delta y)\,\Delta y, \tag{9.22}$$

где $\varepsilon(\Delta y)$ — непрерывная в нуле функция и

$$\lim_{\Delta y \to 0} \varepsilon(\Delta y) = 0.$$

Разделив обе части равенства (9.22) на $\Delta x \neq 0$, получим

$$\frac{\Delta z}{\Delta x} = F'(y_0)\frac{\Delta y}{\Delta x} + \varepsilon(\Delta y)\frac{\Delta y}{\Delta x}. \tag{9.23}$$

Функция $y = f(x)$ имеет производную в точке x_0, т. е. существует предел

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(x_0). \tag{9.24}$$

Из существования производной $f'(x_0)$ следует непрерывность функции $y = f(x)$ в точке x_0:

$$\lim_{\Delta x \to 0} \Delta y = 0.$$

При $\Delta x = 0$ имеем $\Delta y = 0$. Следовательно, приращение Δy, рассматриваемое как функция Δx, непрерывно в точке $\Delta x = 0$. Поэтому, согласно правилу замены переменных в предельных соотношениях, содержащих непрерывные функции (см. п. 5.16),

$$\lim_{\Delta x \to 0} \varepsilon(\Delta y) = 0. \tag{9.25}$$

Теперь из (9.23), переходя к пределу при $\Delta x \to 0$ в силу (9.24) и (9.25), получим формулу (9.21). □

З а м е ч а н и е 1. Формула (9.21) для производной сложной функции справедлива и в том случае, когда под производными понимаются соответствующие односторонние производные, если только предварительно потребовать, чтобы сложная функция, которая необходима для определения рассматриваемой односторонней (или двусторонней) производной, стоящей в левой части формулы (9.21), имела смысл.

СЛЕДСТВИЕ (инвариантность формы первого дифференциала относительно преобразования независимой переменной):

$$dz = F'(y_0)dy = \Phi'(x_0)dx. \tag{9.26}$$

В этой формуле $dy = f'(x)dx$ является дифференциалом функции, а dx — дифференциалом независимой переменной.

Таким образом, дифференциал функции имеет один и тот же вид: произведение производной по некоторой переменной на «дифференциал этой переменной» — независимо от того, является эта переменная, в свою очередь, функцией или независимой переменной.

Докажем это. Согласно формуле (9.6), $dz = \Phi'(x_0)dx$, отсюда, применив формулу (9.21) для производной сложной функции, получим $dz = F'(y_0)f'(x_0)dx$, но $f'(x_0)dx = dy$, поэтому $dz = F'(y_0)dy$. □

Формулу (9.26) можно интерпретировать и несколько иначе, если вспомнить, что дифференциалом функции в точке является функция, линейная относительно дифференциала независимой переменной. Согласно (9.21), дифференциал функции $\Phi(x) = F(f(x))$ имеет вид $d\Phi = F'(y)f'(x_0)dx$, т. е. является результатом подстановки линейной функции $dy = f'(x_0)dx$, с помощью которой задан дифференциал df (где $y = f(x)$), в линейную функцию $dz = F'(y_0)dy$, задающую дифференциал dF (где $z = F(y)$). Иначе говоря, дифференциал композиции $\Phi = F\circ f$ является композицией дифференциалов dF и df:

$$d(F\circ f) = dF\circ df.$$

Отметим, что теорема 5 по индукции распространяется на суперпозицию любого конечного числа функций. Например, для сложной функции вида $z(y(x(y)))$ в случае дифференцируемости функций $z(y)$, $y(x)$ и $x(t)$ в соответствующих точках имеет место формула

$$\frac{dz}{dt} = \frac{dz}{dy}\frac{dy}{dx}\frac{dx}{dt}.$$

Для обозначения производной z сложной функции $z = z(y)$, $y = y(x)$ употребляют также нижний индекс x или y, указывающий, по какой из переменных берется производная, т. е. пишут z'_x или z'_y. Часто для простоты штрих опускают, т. е. вместо z'_x пишут просто z_x. В этих обозначениях формула (9.21) имеет вид $z_x = z_y y_x$.

Примеры. 1. Пусть $y = x^\alpha$, $x > 0$, найдем $\frac{dy}{dx}$. Имеем $x^\alpha = e^u$, где $u = a \ln x$. Замечая, что $\frac{du}{dx} = \frac{\alpha}{x}$, получаем

$$\frac{dx^\alpha}{dx} = \frac{de^u}{dx} = \frac{de^u}{du}\frac{du}{dx} = e^u\frac{\alpha}{x} = e^{\alpha \ln x}\frac{\alpha}{x} = \alpha x^{\alpha - 1}.$$

Следовательно,

$$(x^{\alpha})' = \alpha x^{\alpha - 1}.$$

Например,

$$\text{если } y = x^2, \qquad \text{то} \quad y' = 2x;$$

$$\text{если } y = \frac{1}{x} = x^{-1}, \qquad \text{то} \quad y' = (-1)x^{-2} = -\frac{1}{x^2};$$

$$\text{если } y = \sqrt{x} = x^{1/2}, \quad \text{то} \quad y' = \frac{1}{2}x^{-1/2} = \frac{1}{2\sqrt{x}}.$$

Если функция $y = x^{\alpha}$ определена при $x < 0$, то при этих значениях x она также имеет производную $y' = \alpha x^{\alpha - 1}$.

2. Пусть $y = \ln|x|$, $x \neq 0$; тогда при $x > 0$ имеем

$$y' = (\ln x)' = \frac{1}{x},$$

а при $x < 0$

$$y' = [\ln(-x)]' = \frac{1}{-x}(-x)' = \frac{1}{x}.$$

Таким образом, для всех $x \neq 0$ справедлива формула

$$(\ln|x|)' = \frac{1}{x}. \qquad (9.27)$$

Отсюда, по правилу дифференцирования сложной функции, для любой функции $u(x)$ в точках x, в которых существует производная $u'(x)$, а $u(x) \neq 0$, имеет место соотношение

$$(\ln|u(x)|)' = \frac{u'(x)}{u(x)}. \qquad (9.28)$$

З а м е ч а н и е 2. Формула (9.27) может быть получена и сразу для всех $x \neq 0$ из формулы дифференцирования сложных функций, если вспомнить, что при $x \neq 0$ справедливо равенство $|x|' = \operatorname{sign} x$ (см. пример 1 в конце п. 9.2). Действительно, положив $u = |x|$, для всех $x \neq 0$ получим

$$\frac{d\ln|x|}{dx} = \frac{d\ln u}{du}\frac{du}{dx} = \frac{1}{u}\operatorname{sign} x = \frac{\operatorname{sign} x}{|x|} = \frac{1}{x}.$$

3. Найдем производную функции

$$y = \frac{1}{2a}\ln\left|\frac{x-a}{x+a}\right|, \; x \neq a, \; x \neq -a.$$

В силу формулы (9.28) имеем

$$y' = \frac{1}{2a}\frac{x+a}{x-a}\left(\frac{x-a}{x+a}\right)' = \frac{1}{2a}\frac{x+a}{x-a}\frac{x+a-(x-a)}{(x+a)^2} = \frac{1}{x^2-a^2}.$$

4. Найдем производную функции $y = \ln|x + \sqrt{x^2+A}|$. Аналогично предыдущему примеру получим

$$y' = \frac{1}{x+\sqrt{x^2+A}}(x+\sqrt{x^2+A})' =$$

$$= \frac{1}{x+\sqrt{x^2+A}}\left(1+\frac{x}{\sqrt{x^2+A}}\right) = \frac{1}{\sqrt{x^2+A}}.$$

5. Пусть $y = \ln^2 \arcsin\frac{1}{x}$, $x > 1$. Найдем производную и дифференциал этой функции:

$$y' = (\ln^2 \arcsin 1/x) = 2\ln\arcsin\frac{1}{x}(\ln\arcsin 1/x)' =$$

$$= 2\ln\arcsin\frac{1}{x}\frac{1}{\arcsin 1/x}(\arcsin 1/x)' =$$

$$= 2\frac{\ln\arcsin 1/x}{\arcsin 1/x}\frac{1}{\sqrt{1-1/x^2}}\left(\frac{1}{x}\right)' = -\frac{2\ln\arcsin 1/x}{|x|\sqrt{x^2-1}\arcsin 1/x}.$$

Отсюда дифференциал находится непосредственно по формуле $dy = y'dx$; однако если бы мы еще не имели готового выражения для производной, то дифференциал можно было бы найти и непосредственно, используя его инвариантность относительно выбора переменных:

$$d(\ln^2 \arcsin 1/x) = 2\ln\arcsin 1/x\, d(\ln\arcsin 1/x) =$$

$$= 2\ln\arcsin\frac{1}{x}\frac{1}{\arcsin 1/x}d(\arcsin 1/x) =$$

$$= \frac{2\ln\arcsin 1/x}{\arcsin 1/x}\frac{1}{\sqrt{1-1/x^2}}d\left(\frac{1}{x}\right) = \frac{-2\ln\arcsin 1/x}{|x|\sqrt{x^2-1}\arcsin 1/x}dx.$$

6. Выведем с помощью теоремы 5 еще одну часто применяемую формулу. Пусть $y = u^v$, где $u = u(x) > 0$, $v = v(x)$. Представим данную функцию в виде $y = e^{v\ln u}$ и вычислим $\frac{dy}{dx}$:

$$\frac{dy}{dx} = \frac{de^{v\ln u}}{dx} = e^{v\ln u}\frac{d}{dx}(v\ln u) = u^v\left(\frac{dv}{dx}\ln u + \frac{v}{u}\frac{du}{dx}\right) =$$

$$= u^v\frac{dv}{dx}\ln u + vu^{v-1}\frac{du}{dx}. \qquad (9.29)$$

Таким образом, производная функция u^v равна сумме двух слагаемых, из которых первое совпадает с производной u^v в предположении, что u — постоянная, а второе — с производной u^v в предположении, что v — постоянная.

С помощью правила дифференцирования сложной функции можно находить и *производные функций, заданных неявно*.

7. Пусть дифференцируемая функция $y = y(x)$ задана неявно уравнением $F(x, y) = 0$ (см. п. 5.2). (Вопрос о том, как установить, что данное уравнение на самом деле определяет некоторую функцию, и является ли она дифференцируемой, будет рассмотрен в дальнейшем.) Дифференцируя тождество $F(x, y(x)) \not\equiv 0$ как сложную функцию, можно вычислить производную $\frac{dy}{dx}$.

В качестве примера вычислим производную неявной функции $y(x)$, задаваемой уравнением $x^2 + y^2 = a^2$. В данном конкретном случае существование такой функции не вызывает сомнения, например это $y = \sqrt{a^2 - x^2}$, а также $y = -\sqrt{a^2 - x^2}$. Продифференцируем уравнение $x^2 + y^2 = a^2$, считая y функцией от x. В результате получим $2x + 2yy' = 0$; отсюда $y' = -\frac{x}{y}$.

С подобными задачами приходится сталкиваться в геометрии. Пусть, например, требуется найти касательную к окружности $x^2 + y^2 = 25$ в точке (3, 4). Угловой коэффициент k касательной равен производной; $k = y'$, и, значит, в данном случае $k = -\frac{x}{y}$. Для рассматриваемой точки $k = -\frac{3}{4}$; поэтому уравнение искомой касательной можно записать в виде $y - 4 = -\frac{3(x-3)}{4}$, т. е. $3x + 4y - 25 = 0$.

Применим метод дифференцирования неявных функций к выводу формул, полученных ранее другим путем.

8. Рассмотрим снова функцию $y = u^v$. Логарифмируя, получаем ее неявное задание $\ln y = v \ln u$. Дифференцируя обе части этого уравнения, имеем $\frac{y'}{y} = v' \ln u + \frac{v}{u} u'$ (выражение

$(\ln y)' = \frac{y'}{y}$ называется *логарифмической производной* функции $y(x)$) или $y' = y(v' \ln u + \frac{v}{u} u')$; подставляя $y = u^v$, приходим снова к формуле (9.29).

Рассмотрим еще функцию $y = \arcsin x$. Она неявно задается уравнением $x = \sin y$. Дифференцируя обе части по x, получаем $1 = y' \cos y$, откуда

$$y' = \frac{1}{\cos y} = \frac{1}{\sqrt{1 - \sin^2 y}} = \frac{1}{\sqrt{1 - x^2}},$$

т. е. то же, что и в п. 9.6.

9. В случае, когда функция задана не одной формулой, а несколькими, производную приходится иногда вычислять непосредственно, исходя из определения производной. Найдем, например, производную функции

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x} & \text{при } x \neq 0, \\ 0 & \text{при } x = 0. \end{cases}$$

При $x \neq 0$ производная существует и вычисляется по формулам дифференцирования: $f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}$. В точке же $x = 0$ производная находится непосредственно, исходя из определения:

$$f'(0) = \lim_{x \to 0} \frac{f(x) - f(0)}{x} = \lim_{x \to 0} x \sin \frac{1}{x} = 0.$$

Таким образом, функция $f(x)$ дифференцируема на всей числовой оси.

З а м е ч а н и е. Используя теорему 5, можно все полученные формулы для производных основных элементарных функций записать в более общем виде: если $u = u(x)$ — дифференцируемая функция, то:

$(\sin u)' = u' \cos u$; $\quad (e^u) = e^u u'$;

$(\cos u)' = -u' \sin u$; $\quad (\ln u)' = \frac{u'}{u}$ $(u > 0)$;

$(\operatorname{tg} u)' = \frac{u'}{\cos^2 u}$; $\quad (\arcsin u)' = \frac{u'}{\sqrt{1 - u^2}}$;

$(\operatorname{ctg} u)' = -\frac{u'}{\sin^2 u}$; $\quad (\arccos u)' = -\frac{u'}{\sqrt{1 - u^2}}$;

$(u^\alpha)' = \alpha u^{\alpha - 1} u'$ $(u > 0)$; $\quad (\operatorname{arctg} u)' = \frac{u'}{1 + u^2}$;

$(a^u)' = a^u u' \ln a$; $\quad (\operatorname{arcctg} u)' = -\frac{u'}{1 + u^2}$.

Из приведенных формул видно (при $u = x$), что производные основных элементарных функций являются элементарными функциями.

Полученные же в совокупности формулы дают возможность вычислить производную и дифференциал любой элементарной функции в случае, если эта производная существует.

Следует иметь в виду, однако, что не всякая элементарная функция имеет производные во всех точках своей области определения. Примером элементарной, дифференцируемой не во всех точках функции является функция $|x| = \sqrt{x^2}$, она, как мы знаем, не имеет производной в точке $x = 0$ (см. п. 9.2).

УПРАЖНЕНИЯ. 6. Можно ли доказать формулу $\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}$ при $dy \neq 0$, просто умножив и разделив $\frac{dz}{dx}$ на dy? Можно или нет доказать формулу $\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}$ при $dx \neq 0$, разделив числитель и знаменатель дроби $\frac{dx}{dy}$ на dx?

7. Выяснить, является ли функция

$$f(x) = \begin{cases} x \sin \frac{1}{x} & \text{при } x \neq 0, \\ 0 & \text{при } x = 0 \end{cases}$$

непрерывной в точке $y = 0$. Имеет ли она производную в этой точке? Будет ли она иметь в ней односторонние производные?

9.8. Гиперболические функции и их производные

Определение 5. *Функции* $\frac{e^x + e^{-x}}{2}$ и $\frac{e^x - e^{-x}}{2}$ *называются соответственно гиперболическим косинусом и гиперболическим синусом и обозначаются* $\operatorname{ch} x$ *и* $\operatorname{sh} x$:

$$\frac{e^x + e^{-x}}{2} = \operatorname{ch} x, \quad \frac{e^x - e^{-x}}{2} = \operatorname{sh} x.$$

Справедлива формула

$$\operatorname{ch}^2 x - \operatorname{sh}^2 x = 1. \qquad (9.29)$$

Действительно,

$$\operatorname{ch}^2 x - \operatorname{sh}^2 x = \left(\frac{e^x + e^{-x}}{2}\right)^2 - \left(\frac{e^x - e^{-x}}{2}\right)^2 =$$

$$= \frac{1}{4}(e^{2x} + 2 + e^{-2x} - e^{2x} + 2 - e^{-2x}) = 1.$$

Справедлива также формула

$$\operatorname{sh} 2x = 2 \operatorname{sh} x \operatorname{ch} x;$$

в самом деле,

$$2 \operatorname{sh} x \operatorname{ch} x = 2\frac{e^x + e^{-x}}{2} \frac{e^x - e^{-x}}{2} = \frac{e^{2x} - e^{-2x}}{2} = \operatorname{sh} 2x.$$

Эти формулы напоминают соотношения между обычными (как их иногда называют, круговыми) синусом и косинусом. Для $\operatorname{sh} x$ и $\operatorname{ch} x$ имеется и ряд других соотношений, аналогичных соответствующим формулам для $\sin x$ и $\cos x$. Этим и объясняется название функций $\operatorname{sh} x$ и $\operatorname{ch} x$. Слово же «гиперболический» связано с тем обстоятельством, что формулы

$$x = a \operatorname{ch} t, \ y = a \operatorname{sh} t \tag{9.30}$$

параметрически задают ветвь гиперболы, подобно тому как формулы

$$x = a \cos t, \ y = a \sin t \tag{9.31}$$

параметрически задают окружность. В самом деле, если возвести в квадрат равенства (9.30), вычесть одно из другого и воспользоваться формулой (9.29), то получим $x^2 - y^2 = a^2$, т. е. уравнение равнобочной гиперболы.

Аналогично, из соотношений (9.31) вытекает, что x и y удовлетворяют уравнению $x^2 + y^2 = a^2$, т. е. уравнению окружности.

Найдем производные гиперболических синуса и косинуса. Замечая, что $(-e^{-x})' = -e^{-x}$, имеем

$$(\operatorname{ch} x)' = \left(\frac{e^x + e^{-x}}{2}\right)' = \frac{e^x - e^{-x}}{2} = \operatorname{sh} x,$$

$$(\operatorname{sh} x)' = \left(\frac{e^x - e^{-x}}{2}\right)' = \frac{e^x + e^{-x}}{2} = \operatorname{ch} x.$$

Таким образом,

$$(\operatorname{ch} x)' = \operatorname{sh} x, \ (\operatorname{sh} x)' = \operatorname{ch} x.$$

Частные $\frac{\text{sh } x}{\text{ch } x}$ и $\frac{\text{ch } x}{\text{sh } x}$, по аналогии с обычными синусами и косинусами, называют соответственно *гиперболическим тангенсом* и *гиперболическим котангенсом* и обозначают

$$\frac{\text{sh } x}{\text{ch } x} = \text{th } x, \quad \frac{\text{ch } x}{\sin x} = \text{cth } x.$$

Функции, обратные гиперболическому синусу sh x и гиперболическому косинусу ch x, обозначаются соответственно Areash x, Areach x (читается: «ареасинус, ареакосинус»; area (лат.) — площадь, мера). Появление здесь площади связано с тем, что обратные гиперболические функции связаны с выражением для площадей секторов гиперболы (см. п. 32.1).

Обратные гиперболические функции Areash x и Areach x выражаются через логарифмы иррациональных функций. Покажем это.

Для нахождения функции, обратной sh x, перепишем равенство $y = \frac{e^x + e^{-x}}{2}$ в виде

$$e^{2x} - 2ye^x - 1 = 0$$

и решим получившееся квадратное относительно e^x уравнение. Отбросив отрицательный корень (e^x не бывает отрицательным), получим

$$e^x = y + \sqrt{y^2 + 1}.$$

Отсюда

$$x = \text{Areash } y = \ln\left(y + \sqrt{y^2 + 1}\right).$$

Функция Areash y однозначно определена на всей числовой оси.

Аналогично для $y = \text{ch } x$ получается квадратное относительно e^x уравнение

$$e^{2x} - 2ye^x + 1 = 0,$$

откуда

$$e^x = y \pm \sqrt{y^2 - 1}$$

и, следовательно,

$$x = \text{Areach } y = \ln\left(y \pm \sqrt{y^2 - 1}\right).$$

Функция Areach y определена для всех $y \geqslant 1$ и двухзначна (кроме значения $y = 1$).

УПРАЖНЕНИЕ 8. Вычислить производные функций th x и cth x. Построить графики функций $y = \operatorname{ch} x$, $y = \operatorname{sh} x$, $y = \operatorname{th} x$, $y = \operatorname{cth} x$. Найти производные их обратных функций.

§ 10.

Производные и дифференциалы высших порядков

10.1. Производные высших порядков

Определение 1. *Пусть функция $f(x)$, определенная на интервале (a, b), имеет в каждой точке $x \in (a, b)$ производную $f'(x)$ и пусть $x_0 \in (a, b)$. Если при $x = x_0$ у производной $f'(x)$ функции $f(x)$ существует производная, то она называется второй производной (или производной второго порядка) функции f и обозначается $f''(x_0)$ или $f^{(2)}(x_0)$.*

Таким образом, $f''(x_0) = [f'(x)']'|_{x = x_0}$ или, если опустить обозначение аргумента, $y'' = (y')'$. Аналогично определяется производная $y^{(n)}$ любого порядка $n = 1, 2, ...$: *если существует производная $y^{(n-1)}$ порядка $n - 1$ (при этом под производной нулевого порядка подразумевается сама функция $y^{(0)} = y$, а под производной первого порядка — y'), то, по определению, $y^{(n)} = [y^{(n-1)}]'$.*

Вспоминая определение производной (см. п. 9.1), определение n-й производной в точке x_0 можно записать в виде предела:

$$f^{(n)}(x_0) = \lim_{\Delta x \to 0} \frac{f^{(n-1)}(x_0 + \Delta x) - f^{(n-1)}(x_0)}{\Delta x}, \; n = 1, 2, \ldots .$$

Отметим, что из предположения, что функция f имеет в точке x_0 производную порядка n, следует, в силу определения последней, что в некоторой окрестности точки x_0 у функции f существует производная порядка $n - 1$, а следовательно, при $n > 1$ и все производные более низкого порядка $k < n - 1$, в частности сама функция определена в некоторой окрестности точки x_0. При этом все производные, порядок которых меньше $n - 1$, непрерывны в указанной окрестности, по-

скольку во всех ее точках они имеют производную (см. теоремы 1 и 2 в п. 9.2).

Все сказанное здесь естественным образом переносится и на так называемые односторонние производные высшего порядка, которые читатель без труда определит самостоятельно.

Определение 2. *Функция называется n раз непрерывно дифференцируемой на некотором промежутке, если во всех точках этого промежутка она имеет непрерывные производные до порядка n включительно* ($n = 0, 1, 2, \ldots$).

При этом на каком-либо конце рассматриваемого промежутка в том случае, когда этот конец принадлежит промежутку, под производными, как обычно, понимаются соответствующие односторонние производные.

Для того чтобы функция была n раз непрерывно дифференцируемой на некотором промежутке, достаточно, чтобы она имела на нем непрерывную производную порядка n. Действительно, согласно определению, существование производной порядка n на рассматриваемом промежутке предполагает существование на нем производной порядка $n - 1$ и, поскольку из существования производной какой-либо функции в некоторой точке следует, как это уже отмечалось выше, непрерывность функции в этой точке, производная порядка $n - 1$ непрерывна на данном промежутке. Аналогично, в случае $n > 1$ доказывается непрерывность на рассматриваемом промежутке производной порядка $n - 2$ и т. д.

Примеры. 1. $y = x^3$, $y' = 3x^2$, $y'' = 6x$, $y^{(3)} = 6$, $y^{(4)} = y^{(5)} = \ldots = 0$.

2. $y = a^x$, $y' = a^x \ln a$, $y'' = a^x \ln^2 a$, $y^{(3)} = a^x \ln^3 a$. Вообще, по индукции легко установить, что $y^{(n)} = a^x \ln^n a$. В частности, $(e^x)^{(n)} = e^x$, $n = 0, 1, \ldots$.

3. $y = \sin x$. Вычисляя последовательно производные, получим $y' = \cos x$, $y'' = -\sin x$, $y^{(3)} = -\cos x$, $y^{(4)} = \sin x$, далее производные повторяются в том же порядке. Чтобы записать полученный результат в виде одной формулы, заметим, что $\cos \alpha = \sin (\alpha + \pi/2)$, поэтому $y' = \cos x = \sin (x + \pi/2)$,

$$y'' = \cos\left(x + \frac{\pi}{2}\right) = \sin\left(x + 2\frac{\pi}{2}\right) \text{ и т. д.}$$

По индукции $(\sin x)^{(n)} = \sin(x + n\pi/2)$ для любого $n = 1, 2, \dots$.

4. $y = \cos x$. Замечая, что $-\sin\alpha = \cos(\alpha + \pi/2)$, аналогично предыдущему примеру получим

$$(\cos x)^{(n)} = \cos\left(x + n\frac{\pi}{2}\right), \quad n = 1, 2, \dots .$$

10.2. Производные высших порядков суммы и произведения функций

ТЕОРЕМА 1. *Пусть функции $y_1 = f_1(x)$ и $y_2 = f_2(x)$ имеют производные n-го порядка в точке x_0; тогда функции $y_1 + y_2 = f_1(x) + f_2(x)$ и $y_1y_2 = f_1(x)f_2(x)$ также имеют производные n-го порядка в точке x_0, причем*

$$(y_1 + y_2)^{(n)} = y_1^{(n)} + y_2^{(n)}, \tag{10.1}$$

$$(y_1y_2)^{(n)} = y_1^{(n)}y_2 + C_n^1 y_1^{(n-1)}y_2^{(1)} + $$
$$+ C_n^2 y_1^{(n-2)}y_2^{(2)} + \dots + y_1y_2^{(n)} = \sum_{k=0}^{n} C_n^k y_1^{(n-k)}y_2^{(k)}, \tag{10.2}$$

где, как обычно, C_n^k — число сочетаний из n элементов по k ($k = 0, 1, 2, \dots, n$).

Формулу (10.2) обычно называют формулой Лейбница[1], ее символически можно записать в следующем удобном для запоминания виде:

$$(y_1y_2)^{(n)} = (y_1 + y_2)^{\{n\}}.$$

Индекс $\{n\}$ означает, что выражение $(y_1 + y_2)^{\{n\}}$ записывается подобно биному Ньютона, т. е. в виде суммы с теми же коэффициентами, что и в биномиальной формуле, только степени функций y_1 и y_2 заменяются их производными соответствующего порядка (см. (10.2)).

Формулы (10.1) и (10.2) доказываются по индукции. При $n = 1$, т. е. для производных первого порядка, они были доказаны в п. 9.5. Пусть теперь эти формулы верны для производных n-го порядка. Докажем их справедливость для производных порядка $n + 1$.

[1] Г. Лейбниц (1664—1716) — немецкий философ и математик.

В случае суммы функций имеем

$$(y_1 + y_2)^{(n+1)} = [(y_1 + y_2)^{(n)}]' =$$

$$= (y_1^{(n)} + y_2^{(n)})' = (y_1^{(n)})' + (y_2^{(n)})' = y_1^{(n+1)} + y_2^{(n+1)}.$$

Формула (10.2) доказана.

В случае произведения функций выкладки несколько сложнее:

$$(y_1 y_2)^{(n+1)} = [(y_1 y_2)^{(n)}]' = \left[\sum_{k=0}^{n} C_n^k y_1^{(n-k)} y_2^{(k)}\right]' =$$

$$= \sum_{k=0}^{n} C_n^k \left[\, y_1^{(n+1-k)} y_2^{(k)} + y_1^{(n-k)} y_2^{(k+1)} \,\right] =$$

$$= \sum_{k=0}^{n} C_n^k y_1^{(n+1-k)} y_2^{(k)} + \sum_{k=0}^{n} C_n^k y_1^{(n-k)} y_2^{(k+1)} = y_1^{(n+1)} y_2^{(0)} +$$

$$+ \sum_{k=0}^{n} C_n^k y_1^{(n+1-k)} y_2^{k} + \sum_{k=0}^{n-1} C_n^k y_1^{(n-k)} y_2^{(k+1)} + y_1^{(0)} y_2^{(n+1)}.$$

Здесь мы воспользовались тем, что $C_n^0 = C_n^n = 1$. Теперь изменим индекс суммирования во второй сумме, положив $k = p - 1$; тогда новый индекс суммирования p будет меняться от 1 до n. После этого в полученных суммах объединим попарно слагаемые, содержащие производные одинаковых порядков. Обозначая общий индекс суммирования через p, будем иметь

$$(y_1 y_2)^{(n+1)} = y_1^{(n+1)} y_2^{(0)} +$$

$$+ \sum_{p=1}^{n} (C_n^p + C_n^{p-1})\, y_1^{(n+1-p)} y_2^{(p)} + y_1^{(0)} y_2^{(n+1)}.$$

Отсюда, заметив, что $C_n^p + C_n^{p-1} = C_{n+1}^p$ и что $C_{n+1}^0 = C_{n+1}^{n+1} = 1$, получим

$$(y_1 y_2)^{(n+1)} =$$

$$= y_1^{(n+1)} y_2^{(0)} + \sum_{p=1}^{n} C_{n+1}^p y_1^{(n+1-p)} y_2^{(p)} + y_1^{(0)} y_2^{(n+1)} =$$

$$= \sum_{p=0}^{n+1} C_{n+1}^p y_1^{(n+1-p)} y_2^{(p)}. \ \square$$

СЛЕДСТВИЕ. *Если c — постоянная, а $y = f(x)$ — функция, имеющая производную n-го порядка в точке x_0, то функция $cf(x)$ также имеет производную порядка n при $x = x_0$, причем*

$$(cy)^{(n)} = cy^{(n)}. \qquad (10.3)$$

Действительно, если в формуле (10.2) положить $y_1 = c$, $y_2 = y$, то получится формула (10.3). Впрочем, она следует очевидным образом и из n-кратного применения формулы (9.18) к функции cy.

Рассмотрим пример. Пусть $y = x^3 \sin x$. Найдем с помощью формулы Лейбница производную $y^{(10)}$.

$$(x^3 \sin x)^{(10)} = x^3 \sin\left(x + 10\frac{\pi}{2}\right) + 10 \cdot 3x^2 \sin\left(x + 9\frac{\pi}{2}\right) +$$

$$+ 10 \cdot 9 \cdot 3x \sin\left(x + 8\frac{\pi}{2}\right) + 10 \cdot 9 \cdot 8 \sin\left(x + 7\frac{\pi}{2}\right) =$$

$$= -x^3 \sin x + 30\,x^2 \cos x + 270\,x \sin x - 720 \cos x.$$

10.3. Производные высших порядков от сложных функций, от обратных функций и от функций, заданных параметрически

Пусть функция $y = y(x)$ имеет вторую производную в точке x_0, а $z = z(y)$ — вторую производную в точке $y_0 = y(x_0)$. Тогда сложная функция $z[y(x)]$ имеет при $x = x_0$ вторую производную, причем

$$z'_{xx} = z''_{yy}\,{y'_x}^2 + z'_y\,y''_{xx}. \qquad (10.4)$$

Действительно, поскольку существуют производные $y''(x_0)$ и $z''(y_0)$, существуют также $y'(x_0)$ и $z'(y_0)$. Следовательно, функции $y(x)$ и $z(y)$ непрерывны соответственно в точках x_0 и y_0. Поэтому в некоторой окрестности точки x_0 определена сложная функция $z = z[y(x)]$. Дифференцируя ее и опуская для простоты обозначение аргумента, имеем $z'_x = z'_y y'_x$; дифференцируя еще раз по x, получим

$$z''_{xx} = (z'_y)'_x\, y'_x + z'_y\, y''_{xx} = z''_{yy}\,{y'_x}^2 + z'_y\, y''_{xx}. \ \square$$

Аналогичным образом вычисляются, при соответствующих предположениях, и *производные высших порядков сложной функции*. Этот метод позволяет также доказывать существование и находить производные высших порядков от обратной функции.

Пусть функция $y = y(x)$ непрерывна и строго монотонна в некоторой окрестности точки x_0 (см. п. 9.6) *и пусть*

при $x = x_0$ *существуют производные* y' *и* y'', *причем* $y'(x_0) \neq 0$; *тогда и обратная функция* $x = x(y)$ *имеет вторую производную в точке* $y_0 = y(x_0)$, *причем она может быть выражена через значения производных* y' *и* y'' *функции* $y(x)$ *при* $x = x_0$.

В самом деле, опуская, как и выше, обозначения аргумента, согласно теореме 4 § 9 (см. п. 9.6), имеем $x'_y = 1/y'_x$. Вычисляя производную по y от обеих частей и применяя к правой части правило дифференцирования сложной функции, получаем

$$x''_{yy} = (x'_y)'_y = \left(\frac{1}{y'_x}\right)'_x x'_y = \frac{y''_{xx}}{y'^2_x} \cdot \frac{1}{y''_x} = -\frac{y''_{xx}}{y'^3}.$$

Аналогично при соответствующих предположениях вычисляют и *производные высших порядков для обратной функции*.

Подобным же образом можно поступать и в случае так называемого параметрического задания функции.

Определение 3. *Пусть функции* $x = x(t)$ *и* $y = y(t)$ *определены в некоторой окрестности точки* t_0 *и одна из них, например* $x = x(t)$, *непрерывна и строго монотонна в указанной окрестности; тогда существует обратная* $x(t)$ *функция* $t = t(x)$, *и в некоторой окрестности точки* $x_0 = x(t_0)$ *имеет смысл композиция* $y(t(x))$. *Эта функция* y *от* x *называется параметрически заданной формулами* $x = x(t)$, $y = y(t)$ *функцией.*

Выведем формулы для дифференцирования параметрически заданных функций.

Если функции $x(t)$ *и* $y(t)$ *имеют в точке* t_0 *производные и если* $x'(t_0) \neq 0$, *то параметрически заданная функция* $y(t(x))$ *также имеет в точке* $x_0 = x(t_0)$ *производную, причем*

$$y'_x = \frac{y'_t(t_0)}{x'_t(t_0)}. \qquad (10.5)$$

В самом деле, по правилу дифференцирования сложной функции имеем (опуская обозначение аргумента)

$$y'_x = y'_t\, t'_x; \qquad (10.6)$$

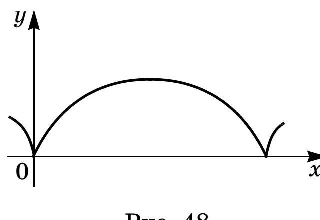


Рис. 48

по правилу же дифференцирования обратной функции

$$t'_x = \frac{1}{x'_t}. \qquad (10.7)$$

Из формул (10.6) и (10.7) и следует формула (10.5).

Если, кроме того, существуют $x''_{tt}(t_0)$ *и* $y''_{tt}(t_0)$, *то существует и* $y''_{xx}(x_0)$, *причем*

$$y''_{xx} = (y'_x)'_x = \left(\frac{y'_t}{x'_t}\right)'_x = \left(\frac{y'_t}{x'_t}\right)'_t t'_x = \frac{y''_{tt}x'_t - y'_t x''_{tt}}{x'^3_t}.$$

Аналогично вычисляются *производные более высокого порядка параметрически заданных функций*.

Рассмотрим в качестве примера параметрически заданную функцию

$$x = a(t - \sin t),\ y = a(1 - \cos t),\ (a \neq 0, -\infty < t < +\infty). \qquad (10.8)$$

Ее график изображен на рисунке 48. Пусть для определенности $a > 0$; тогда функция $x(t) = a(t - \sin t)$ строго возрастает. Действительно, пусть $\Delta t > 0$; тогда, замечая, что $0 < \sin\frac{\Delta t}{2} < \frac{\Delta t}{2}$, имеем

$$x(t + \Delta t) - x(t) = a\{\Delta t - [\sin(t + \Delta t) - \sin t]\} =$$

$$= a\left[\Delta t - 2\cos\left(t + \frac{\Delta t}{2}\right)\sin\frac{\Delta t}{2}\right] > a\left(\Delta t - 2 \cdot 1 \cdot \frac{\Delta t}{2}\right) = 0,$$

что и означает строго монотонное возрастание функции $x(t)$. В силу этого, существует однозначная обратная функция $t = t(x)$. Далее, $x'_t = a(1 - \cos t) = 2a\sin^2\frac{t}{2} \geqslant 0$, $y'_t = a\sin t$ и x'_t обращается в нуль только в точках вида $t = 2k\pi$, $k = 0, \pm 1, \pm 2, \ldots$. Поэтому если $t \neq 2k\pi$, то, согласно правилу дифференцирования функции, заданной параметрически, имеем

$$y'_x = \frac{y'_x}{x'_t} = \frac{\sin t}{2\sin^2\frac{t}{2}} = \operatorname{ctg}\frac{t}{2};$$

$$y''_{xx} = \left(\operatorname{ctg}\frac{t}{2}\right)'_x = \left(\operatorname{ctg}\right)'_t\frac{1}{x'_t} = -\frac{1}{2\sin^2\frac{t}{2}}\ \frac{1}{2a\sin^2\frac{t}{2}} = -\frac{1}{4a\sin^4\frac{t}{2}}.$$

10.4. Дифференциалы высших порядков

В настоящем пункте для удобства будем иногда вместо символа дифференцирования d писать букву δ, т. е. вместо dy, dx писать δy, δx.

Пусть функция $y = f(x)$ дифференцируема на некотором интервале (a, b). Как известно, ее дифференциал

$$dy = f'(x)dx,$$

который называется также ее первым дифференциалом, зависит от двух переменных: x и dx. Пусть функция $f'(x)$, в свою очередь, дифференцируема в некоторой точке $x_0 \in (a, b)$. Тогда дифференциал в этой точке функции dy, рассматриваемой как функция только от x (т. е. при некотором фиксированном dx), если для его обозначения использовать символ δ, имеет вид

$$\delta(dy) = \delta[f'(x)dx]\big|_{x = x_0} = [f(x)dx]'\big|_{x = x_0}\ \delta x = f''(x_0)dx\delta x.$$

Определение 4. *Значение дифференциала $\delta(dy)$, т. е. дифференциала от первого дифференциала, в некоторой точке x_0 при $dx = \delta x$ называется вторым дифференциалом функции f в этой точке и обозначается через d^2y, т. е.*

$$d^2y = f''(x_0)dx^2. \tag{10.9}$$

(*Через dx^2 и вообще через dx^n, $n \in \mathbf{N}$, обозначается $(dx)^2$, соответственно $(dx)^n$, а не $d(x^n)$.*)

Заметим, что, в силу этого определения, $d^2x = 0$, так как при вычислении дифференциалов мы считаем приращением $dx = \Delta x$ постоянным.

Подобным же образом в том случае, когда производная $(n - 1)$-го порядка $y^{(n-1)}$, дифференцируема в точке x_0 или, что эквивалентно, когда при $x = x_0$ существует производная n-го порядка $y^{(n)}$, определяется *дифференциал n-го порядка* d^ny функции $y = f(x)$ в точке x_0 как дифференциал $\delta(d^{n-1}y)$ от дифференциала $(n - 1)$-го порядка $d^{n-1}y$, в котором $\delta x = dx$:

$$d^ny = \delta(d^{n-1}y)|_{\delta x = dx}.$$

Покажем, что справедлива формула

$$d^ny = y^n dx^n,\ n = 1, 2, \ldots . \tag{10.10}$$

Ее доказательство проведем по индукции. Для $n = 1$ и $n = 2$ она доказана. Пусть эта формула верна для дифференциалов порядка $n - 1$:

$$d^{n-1}y = y^{(n-1)}dx^{n-1}.$$

Тогда, согласно данному выше определению, для вычисления дифференциала n-го порядка $d^{(n)}y$ необходимо вычислить сначала дифференциал (мы его обозначим символом δ) от $d^{n-1}y$:

$$\delta(d^{n-1}y) = \delta(y^{(n-1)}dx^{(n-1)}) = (y^{(n-1)}dx^{n-1})'\delta x = y^{(n)}\delta x dx^{n-1},$$

а затем положить $\delta x = dx$:

$$d^n y = \delta(d^{n-1}y)|_{\delta x = dx} = y^{(n)}dx^n. \ \square$$

Из формулы (10.10) следует, что

$$y^{(n)} = \frac{d^n y}{dx^n}. \qquad (10.11)$$

Приведем некоторые **свойства дифференциалов высших порядков:**

1^0. $d^n(y_1 + y_2) = d^n y_1 + d^n y_2$.

2^0. $d^n(cy) = cd^n y$, c — постоянная.

3^0. $d^n(y_1 y_2) = \sum\limits_{k=0}^{n} C_n^k dy_1^{n-k} dy_2^k$, *или, если использовать символическую запись,*

$$d^n(y_1 y_2) = (dy_1 + dy_2)^{\{n\}},$$

где выражение $(dy_1 + dy_2)^{\{n\}}$ *записывается по аналогии с биномиальной формулой Ньютона, т. е. является суммой вида*

$$C_n^k d^{n-k} y_1 d^k y_2;$$

при этом для любой функции и считается, что $d^0 u = u^{(0)}dx^0 = u$.

Эти свойства непосредственно следуют из соответствующих формул для производных n-го порядка (см. формулы (10.1), (10.2), (10.3) и (10.10)).

З а м е ч а н и е. Формулы (10.10) и (10.11) справедливы, вообще говоря, при $n > 1$ (в отличие от случая $n = 1$) только тогда, когда x является независимым переменным. В случае дифференциалов высших порядков по зависимым переменным дело обстоит сложнее.

Пусть $z = z(y)$, $y = y(x)$, имеет смысл суперпозиция $z[y(x)]$ и функции $z(y)$ и $y(x)$ дважды дифференцируемы. Тогда $dz = z'_y\, dy$. Дифференцируя еще раз и не прибегая для простоты к символу δ, т. е. считая запись $d(dz)$ равносильной записи $\delta(dz)|_{\delta x = dx}$ (так всегда и поступают на практике), причем здесь под $\delta(dz)$ понимается дифференциал по x от функции $dz = z'_y(y)dy = z'_y\,[y(x)]y'_x\,(x)dx$, получаем

$$d^2z = d(dz) = d(z'_y dy) = d(z'_y\,)dy + z'_y d(dy) = \\ = z''_{yy} dy^2 + z'_y d^2y \qquad (10.12)$$

(мы написали $dz'_y = z''_{yy}dy$ на основании формулы (9.26), т. е. использовав инвариантность первого дифференциала).

Сравнивая формулы (10.9) и (10.12), видим, что они отличаются вторым членом, и так как, вообще говоря, $d^2y \neq 0$, то существенно различны. Разделив обе части равенства (10.12) на dx^2, получим формулу второй производной для сложной функции

$$z''_{xx} = z''_{yy}\, y'^2_x + z'_y\, y''_{xx}\,,$$

которая была получена ранее (см. (10.4)) другим путем.

Подобным же образом могут быть вычислены дифференциалы и производные высших порядков сложной функции.

§ 11.

Теоремы о среднем для дифференцируемых функций

11.1. Теорема Ферма

В терминах производных оказывается удобным описывать различные свойства функций. Прежде всего укажем характеристическое свойство точек, в которых функция принимает наибольшее или наименьшее значение. Напомним, что если функция f определена на некотором множестве X, то говорят, что она принимает в точке x_0 наибольшее (наименьшее) значение на множестве X, если для всех точек $x \in X$ выполняется неравенство $f(x) \leqslant f(x_0)$ (неравенство $f(x) \geqslant f(x_0)$).

Если для всех $x \in X$ и $x \neq x_0$ выполняется неравенство $f(x) < f(x_0)$ (неравенство $f(x) > f(x_0)$), то говорят, что в точке x_0

функция f принимает строго наибольшее (наименьшее) значение на множестве X.

Точки, в которых функция принимает значения (строгого) максимума или минимума, называются *точками* (строгого) *экстремума*.

ТЕОРЕМА 1 (Ферма[1]). *Если функция определена в некоторой окрестности точки, принимает в этой точке наибольшее (наименьшее) значение и имеет в ней конечную или определенного знака бесконечную производную, то эта производная равна нулю.*

Доказательство. Пусть функция f определена в окрестности $U(x_0)$ точки x_0 и принимает для определенности при $x = x_0$ наибольшее значение, т. е. для всех $x \in U(x_0)$ выполняется неравенство $f(x) \leqslant f(x_0)$. Тогда если $x < x_0$,

$$\frac{f(x) - f(x_0)}{x - x_0} \geqslant 0, \tag{11.1}$$

а если $x > x_0$, то

$$\frac{f(x) - f(x_0)}{x - x_0} \leqslant 0. \tag{11.2}$$

По условию теоремы существует конечный или определенного знака бесконечный предел

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0),$$

поэтому в неравенствах (11.1) и (11.2) можно перейти к пределу при $x \to x_0$ (см. свойство 4^0 пределов функций в п. 5.10). В результате получим соответственно $f'(x_0) \geqslant 0$ и $f'(x_0) \leqslant 0$. Следовательно, $f'(x_0) = 0$. □

Замечание 1. Формулировка теоремы Ферма с первого взгляда может показаться неестественной: в предположениях говорится о существовании бесконечных производных, а в утверждении — о равенстве нулю производной. Однако на самом деле формулировка теоремы вполне корректна: a priori предполагается, что в точке существует производная (конечная или определенного знака бесконечная) и доказывается, что при выполнении дополнительного условия о достижении в рассматриваемой точке наибольшего (наименьшего) значения указанная производная равна нулю. Иначе говоря, доказывается, что в точке, в которой

[1] П. Ферма (1601—1665) — французский математик.

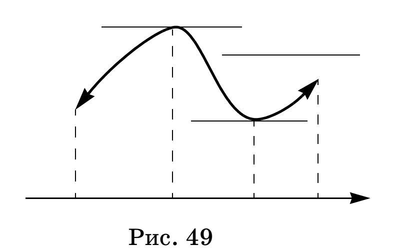


Рис. 49

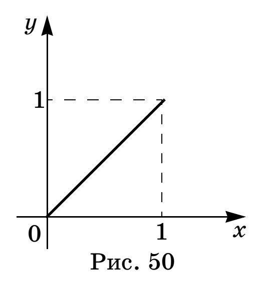


Рис. 50

принимается наибольшее или наименьшее в некоторой ее окрестности значение функции, не может существовать ни конечная, ни равная нулю производная функции, ни определенного знака бесконечная производная. Поэтому в точке, в которой достигается наибольшее или наименьшее в ее окрестности значение функции, возможны следующие случаи: в этой точке существует конечная, равная нулю производная, существует знаконеопределенная бесконечная производная, не существует никакой производной ни конечной, ни бесконечной. Примером функции, для которой осуществляется при $x = 0$ первый случай, является функция $f_1(x) = x^2$, второй случай: $f_2(x) = \sqrt[3]{x^2}$, третий: $f_3(x) = |x|$. Все эти функции принимают при $x = 0$ наименьшее значение, также равное нулю, производная первой из них равна нулю: $f_1'(0) = 0$, второй бесконечности неопределенного знака: $f_2'(0) = \infty$, а у третьей производная в точке $x = 0$ не существует.

Геометрическая интерпретация теоремы Ферма состоит в том, что если при $x = x_0$ дифференцируемая функция f принимает наибольшее (наименьшее) значение на некоторой окрестности точки x_0, то касательная к графику функции в точке $(x_0, f(x_0))$ параллельна оси Ox (рис. 49).

З а м е ч а н и е 2. Если функция f принимает наибольшее (наименьшее) значение при $x = x_0$ по сравнению с ее значениями в точках, лежащих по одну сторону от точки x_0, и имеет в x_0 соответствующую одностороннюю производную, то эта производная может быть не равна нулю. Так, например, функция $f(x) = x$, рассматриваемая на отрезке $[0, 1]$, принимает при $x = 0$ минимальное, а при $x = 1$ — максимальное значение, однако как в той, так и в другой точке производная равна единице (рис. 50).

11.2. Теоремы Ролля, Лагранжа и Коши о средних значениях

ТЕОРЕМА 2 (теорема Ролля[1]). *Пусть функция* f:

1) *непрерывна на отрезке* $[a, b]$;

2) *имеет в каждой точке интервала* (a, b) *конечную или определенного знака бесконечную производную*;

3) *принимает равные значения на концах отрезка, т. е.* $f(a) = f(b)$.

Тогда существует хотя бы одна такая точка ξ, $a < \xi < b$, *что* $f'(\xi) = 0$.

Геометрический смысл теоремы Ролля состоит в том, что на графике функции, удовлетворяющей условиям теоремы Ролля, имеется по крайней мере одна точка, в которой касательная горизонтальна (рис. 51).

Доказательство. Если для любой точки x интервала (a, b) выполняется равенство $f(x) = f(a) = f(b)$, то функция f является постоянной на этом интервале и поэтому для любой точки $\xi \in (a, b)$ выполняется условие $f'(\xi) = 0$.

Пусть существует точка $x_0 \in (a, b)$, для которой $f(x_0) \neq f(a)$, например, $f(x_0) > f(a)$. Согласно теореме Вейерштрасса о достижимости непрерывной на отрезке функцией своих наибольшего и наименьшего значений (см. теорему 1 в п. 6.1), существует такая точка $\xi \in [a, b]$, в которой функция f принимает наибольшее значение. Тогда

$$f(\xi) \geqslant f(x_0) > f(a) = f(b).$$

Поэтому $\xi \neq a$ и $\xi \neq b$, т. е. точка ξ принадлежит интервалу (a, b) и функция f принимает в ней наибольшее значение. Следовательно, согласно теореме Ферма (см. теорему 1 в п. 11.1) выполняется равенство $f'(\xi) = 0$. □

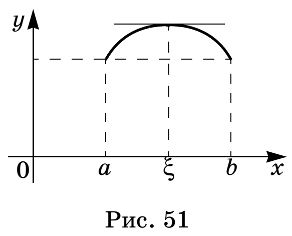


Рис. 51

Заметим, что все предпосылки теоремы Ролля существенны. Чтобы в этом убедиться, достаточно привести примеры функций, для которых выполнялись бы два из трех условий теоремы, третье же не выполнялось бы и у которых не существует точки ξ

[1] М. Ролль (1652—1719) — французский математик.

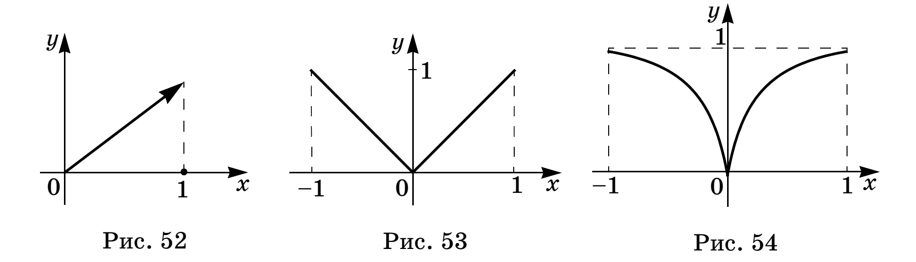


Рис. 52 Рис. 53 Рис. 54

такой, что $f'(\xi) = 0$. (При этом, в силу условия 3, в котором говорится о значениях функции в концевых точках промежутка, следует рассматривать лишь функции, определенные на отрезках.)

Функция $f(x)$, определенная на отрезке $[0, 1]$ и равная x, если $0 \leqslant x < 1$, и 0, если $x = 1$, удовлетворяет условиям 2 и 3, но не удовлетворяет условию 1 (рис. 52).

Функция $f(x) = |x|$, $x \in [-1; 1]$ удовлетворяет условиям 1 и 3, но не удовлетворяет условию 2 (рис. 53).

Наконец, функция $f(x) = x$, $x \in [0; 1]$ удовлетворяет условиям 1 и 2, но не удовлетворяет условию 3 (см. рис. 50).

Для всех этих функций не существует точки, в которой их производная обращалась бы в нуль.

Обратим внимание на то, что по условиям теоремы Ролля отрезок $[a, b]$ может содержать точки, в которых функция имеет определенного знака бесконечную производную, т. е. в которых $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = +\infty$ или $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = -\infty$. Это требование нельзя ослабить, заменив его условием $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \infty$. Например, для функции $f(x) = \sqrt{|x|}$, $-1 \leqslant x \leqslant 1$, не существует точки $\xi \in [-1, +1]$, в которой производная этой функции обращалась бы в нуль. Вместе с тем функция $f(x) = \sqrt{|x|}$ удовлетворяет всем условиям теоремы Ролля на отрезке $[-1, 1]$, за исключением того, что в точке $x = 0$ эта функция не имеет ни конечной, ни определенного знака бесконечной производной (рис. 54).

В самом деле, для этой точки $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \infty$, причем этот предел не является бесконечностью определенного знака.

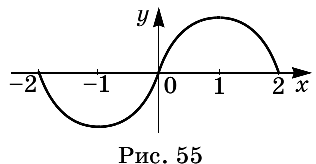


Рис. 55

Примером функции, удовлетворяющей теореме Ролля и имеющей в некоторой точке определенного знака бесконечную производную, является функция

$$f(x) = \begin{cases} \sqrt{1 - (x - 1)^2}, 0 \leqslant x \leqslant 2, \\ -\sqrt{1 - (x + 1)^2}, -2 \leqslant x \leqslant 0. \end{cases}$$

Эта функция непрерывна на отрезке $[-2, 2]$, дифференцируема во всех точках интервала $(-2, 2)$, кроме точки $x = 0$, в которой $f'(0) = +\infty$ и $f(-2) = f(2)$ (рис. 55). В согласии с теоремой Ролля у нее имеется точка (даже две), в которой производная равна нулю ($x = \pm 1$). Графиком этой функции являются две полуокружности, сопряженные в точке $(0; 0)$.

Этот пример показывает целесообразность рассмотрения в данном случае не только функций, имеющих конечные производные, но и функций с определенного знака бесконечными производными.

Заметим, что построением соответствующих примеров (если, конечно, это удается сделать) и проверяют обычно в математике существенность тех или иных условий доказываемых теорем.

В дальнейшем мы не будем проверять необходимость условий теорем, предоставляя это делать читателю по мере внутренней потребности.

Из теоремы Ролля следует, что если функция непрерывна на некотором отрезке, обращается в нуль на его концах и дифференцируема во всех его внутренних точках, то существует его внутренняя точка, в которой производная обращается в нуль. Короче говоря, *между двумя нулями дифференцируемой функции всегда лежит хотя бы один нуль ее производной.*

УПРАЖНЕНИЯ. 1. Доказать, что если функция f удовлетворяет условиям теоремы Ролля на отрезке $[a, b]$ и не является постоянной, то на этом отрезке существуют такие точки ξ_1 и ξ_2, что $f'(\xi_1) > 0$ и $f'(\xi_2) < 0$.

2. Привести пример функции, непрерывной на отрезке $[a, b]$, имеющей производную в каждой точке интервала (a, b), но не имеющей производной (односторонней) в точке a.

ТЕОРЕМА 3 (теорема Лагранжа[1]). *Если функция f непрерывна на отрезке* $[a, b]$ *и в каждой точке интервала* (a, b)

[1] Ж.-Л. Лагранж (1736—1813) — французский математик и механик.

имеет конечную или определенного знака бесконечную производную, то в этом интервале существует по крайней мере одна такая точка ξ, *что*

$$f(b) - f(a) = f'(\xi)(b - a). \quad (11.3)$$

Эта теорема является, очевидно, обобщением теоремы Ролля.

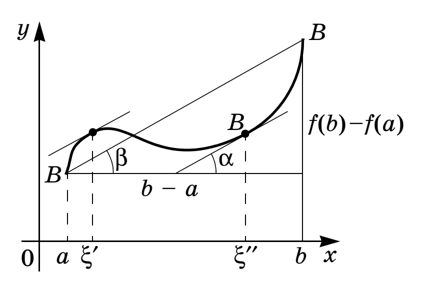


Рис. 56

Доказательство. Рассмотрим вспомогательную функцию

$$F(x) = f(x) - \lambda x \quad (11.4)$$

и определим число λ так, чтобы $F(a) = F(b)$, т. е. чтобы $f(a) - \lambda a = f(b) - \lambda b$. Это равносильно тому, что

$$\lambda = \frac{f(b) - f(a)}{b - a}. \quad (11.5)$$

Для функции F выполняются все условия теоремы Ролля. Действительно, функция $f(x)$ непрерывна на отрезке $[a, b]$, а функция λx, будучи линейной, непрерывна на всей числовой оси; поэтому и функция $F(x) = f(x) - \lambda x$ также непрерывна на отрезке $[a, b]$. Функция f имеет в каждой точке интервала (a, b) конечную или определенного знака бесконечную производную, а функция λx — конечную производную во всех точках числовой оси, поэтому их разность $F(x)$ также имеет всюду в интервале (a, b) конечную или определенного знака бесконечную производную (см. замечание в п. 9.5). Наконец, на концах отрезка $[a, b]$, в силу выбора λ (см. (11.5)), функция F принимает одинаковые значения. Поэтому существует хотя бы одна такая точка ξ $(a < \xi < b)$, что $F'(\xi) = 0$. Из (11.4) получаем $F'(x) = f'(x) - \lambda$, поэтому $f'(\xi) - \lambda = 0$. Подставив сюда λ из (11.5), получим

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}. \ \square \quad (11.6)$$

Геометрический смысл теоремы Лагранжа состоит в следующем. Пусть $A = (a, f(a))$, $B = (b, f(b))$ — концы графика функции f, AB — хорда, соединяющая точки A и B (рис. 56). Тогда отношение $\frac{f(b) - f(a)}{b - a}$ равно тангенсу угла β между хордой AB и осью Ox, т. е.

$$\frac{f(b) - f(a)}{b - a} = \operatorname{tg} \beta,$$

а производная $f'(\xi)$, как известно (см. п. 9.3), равна тангенсу угла α между касательной к графику функции f в точке $(\xi, f(\xi))$ и положительным направлением оси Ox, т. е. $f'(\xi) = \operatorname{tg}\alpha$. Поэтому равенство (11.6) может быть переписано в виде $\operatorname{tg}\alpha = \operatorname{tg}\beta$. Таким образом, теорема Лагранжа показывает, что в интервале (a, b) должна найтись точка ξ (может быть, и не одна: на рисунке 56 условию теоремы удовлетворяют точки ξ' и ξ''), в которой касательная к графику параллельна хорде AB.

Теорема Лагранжа найдет ряд важных приложений в дальнейшем.

Приведем другие формы записи формулы (11.3). Пусть $a < \xi < b$ и $\frac{\xi - a}{b - a} = \theta$. Тогда

$$\xi = a + \theta(b - a),\ 0 < \theta < 1. \tag{11.7}$$

Наоборот, если ξ выражается формулой (11.7), то, как легко видеть, $a < \xi < b$. Таким образом, в виде (11.7) могут быть представлены все точки интервала (a, b) и только они. Поэтому формула (11.3) может быть записана в виде

$$f(b) - f(a) = f'(a + \theta(b - a))(b - a),\ \ 0 < \theta < 1. \tag{11.8}$$

Положим теперь $a = x,\ b - a = \Delta x$ и, значит, $b = x + \Delta x$; тогда (11.8) можно переписать в виде

$$f(x + \Delta x) - f(x) = f'(x + \theta\,\Delta x)\,\Delta x,\ \ 0 < \theta < 1. \tag{11.9}$$

Формулу (11.9), а также каждую из равнозначных ей формул (11.3) и (11.8) называют *формулой конечных приращений Лагранжа* или просто *формулой конечных приращений* в отличие от приближенного равенства

$$f(x + \Delta x) - f(x) \approx f'(x)\,\Delta x, \tag{11.10}$$

которое иногда называют *формулой бесконечно малых приращений*. Она выражает тот факт, что левая и правая части приближенного равенства (11.10) равны между собой для дифференцируемой в точке x функции f «с точностью до бесконечно малых более высокого порядка, чем приращение Δx при $\Delta x \to 0$».

З а м е ч а н и е. Формула Лагранжа (11.3) может быть представлена в виде

$$f(a) - f(b) = f'(\xi)(a - b),$$

где $a < b$. Отсюда следует, что формула (11.3) справедлива не только при $a < b$, но и для $a > b$, причем и в этом случае $\xi = a + \theta(b - a)$, $0 < \theta < 1$ (здесь $b - a < 0$, $\xi - a < 0$).

Отметим три следствия из теоремы Лагранжа, полезные для дальнейшего.

СЛЕДСТВИЕ 1. *Если функция f непрерывна на некотором промежутке* (*конечном или бесконечном*) *и во всех его внутренних точках имеет производную, равную нулю, то функция постоянна на этом промежутке.*

Доказательство. Пусть функция f непрерывна на промежутке с концами a и b, $-\infty \leqslant a < b + \infty$, и дифференцируема в его внутренних точках. Выберем на этом промежутке произвольно точки x_1 и x_2 так, что $x_1 < x_2$; тогда, очевидно, функция f является непрерывной на отрезке $[x_1, x_2]$ и дифференцируемой на интервале (x_1, x_2). Поэтому, по теореме Лагранжа,

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1), \quad x_1 < \xi < x_2. \qquad (11.11)$$

По условию $f'(x) = 0$ на (x_1, x_2), в частности $f'(\xi) = 0$, так как $\xi \in (x_1, x_2)$. Таким образом, из формулы (11.11) следует, что $f(x_1) = f(x_2)$, а поскольку x_1 и x_2 — произвольные точки рассматриваемого промежутка, то это и означает, что функция f постоянна на этом промежутке. □

Следствие 1 имеет наглядную механическую интерпретацию: если функция $y = f(x)$ является законом движения материальной точки по прямой, x — время, y — расстояние (со знаком) от начала отсчета на прямой, то условие $f'(x) = 0$ для всех $x \in (a, b)$ означает, что скорость рассматриваемой точки в течение временно́го интервала (a, b) все время равна нулю, т. е. точка неподвижна, но тогда за это время положение точки, а потому и пройденный ею путь не изменятся. Это и означает, что функция $f(x)$ постоянна на интервале (a, b).

СЛЕДСТВИЕ 2. *Если функция f и g непрерывны на некотором промежутке и во всех его внутренних точках имеют равные производные*

$$f'(x) = g'(x),$$

то эти функции отличаются на рассматриваемом промежутке лишь на постоянную:

$$f(x) = g(x) + C, \qquad (11.12)$$

где C — константа.

Действительно, функция $F = f - g$ удовлетворяет условиям следствия 1, т. е. F непрерывна на заданном промежутке и $F' = 0$ во всех внутренних его точках. Поэтому $F = C$, т. е. имеет место равенство (11.12). □

СЛЕДСТВИЕ 3. *Пусть функция* $\varphi(x)$:
1) *непрерывна на полуинтервале* $[a, b)$;
2) *дифференцируема во всех точках интервала* (a, b);
3) *у производной* $\varphi'(x)$ *существует конечный предел справа в точке* $x = a$.

Тогда в этой точке у функции $\varphi(x)$ *существует правосторонняя производная* $\varphi'_+(a)$ *и*

$$\varphi'_+(a) = \lim_{x \to a+0} \varphi'(x).$$

Пусть в точке $x = a$ у производной $\varphi'(x)$ существует конечный предел справа и $\lim\limits_{x \to a+0} \varphi'(x) = A$. Применив к функции $\varphi(x)$ теорему Лагранжа на отрезке $[a, x]$, где $a < x < b$, будем иметь

$$\varphi(x) - \varphi(a) = \varphi'(\xi)(x - a),\ a < \xi < x,$$

откуда

$$\frac{\varphi(x) - \varphi(a)}{x - a} = \varphi'(\xi).$$

Точка ξ является функцией точки x и притом, вообще говоря, многозначной: $\xi = \xi(x)$. Выберем произвольно для каждого $x \in (a, b)$ одно какое-либо значение ξ; тогда получим однозначную функцию $\xi(x)$ (как говорят, однозначную ветвь многозначной функции). Так как $a < \xi(x) < x$, то

$$\lim_{x \to a+0} \xi(x) = a.$$

Применив правило замены переменного для пределов функций (см. п. 5.16), получим, что существует предел

$$\lim_{x \to a+0} \varphi'(\xi) = A,$$

следовательно, существует и предел

$$\lim_{x \to a+0} \frac{\varphi(x) - \varphi(a)}{x - a} = A.$$

Это и означает, что правосторонняя производная $\varphi'_+(a)$ существует и равна A. □

Утверждение, аналогичное следствию 4, имеет место и для левосторонних производных, откуда следует, что функция, являющаяся производной некоторой функции, не может иметь устранимых точек разрыва.

УПРАЖНЕНИЕ 3. Пусть функция f непрерывна на интервале $(a,\ b)$ и дифференцируема во всех точках этого интервала, кроме, быть может, некоторой точки $x_0 \in (a,\ b)$. Пусть существуют $\lim\limits_{x \to x_0 - 0} f'(x)$ и $\lim\limits_{x \to x_0 + 0} f'(x)$, причем они не равны между собой. Доказать, что при этих предположениях производная $f'(x_0)$ не существует.

В качестве непосредственного применения теоремы Лагранжа получим с ее помощью достаточное условие равномерной непрерывности дифференцируемых функций.

ЛЕММА 1. *Если функция $f(x)$ определена и имеет ограниченную производную на некотором интервале $(a,\ b)$, то она равномерно непрерывна на этом интервале.*

Действительно, если $|f'(x)| \leqslant c$ (c — постоянная) на $(a,\ b)$, то с помощью формулы конечных приращений Лагранжа (см. п. 11.2) получим

$$|f(x') - f(x)| = |f'(\xi)(x' - x)| \leqslant c|x' - x|,$$
$$a < x < b,\ a < x' < b,\ a < \xi < b. \tag{19.23}$$

Поэтому для $\varepsilon > 0$ достаточно взять $\delta = \varepsilon/c$; тогда если $|x' - x| < \delta$, $a < x < b$, $a < x' < b$, то, в силу (19.23), справедливо неравенство $|f(x') - f(x)| < \varepsilon$, что и означает равномерную непрерывность функции f на $(a,\ b)$. □

Аналогичный результат имеет место для любого промежутка, конечного или бесконечного.

В теоремах Ролля и Лагранжа (а также и в следующей ниже теореме Коши) речь идет о существовании некоторой точки ξ, $a < \xi < b$; ее можно назвать «средней точкой», для которой выполняется то или иное равенство. Этим и объясняется название «теоремы о среднем» для этой группы теорем. Докажем последнее нужное нам утверждение этого типа.

ТЕОРЕМА 4 (теорема Коши). *Пусть функции f и g:*

1) *непрерывны на отрезке* $[a,\ b]$;
2) *имеют производные в каждой точке интервала* $(a,\ b)$;
3) $g' \neq 0$ *во всех точках интервала* $(a,\ b)$.

Тогда существует такая точка ξ, $a < \xi < b$, *что*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}. \qquad (11.13)$$

Заметим, что из условий теоремы следует, что формула (11.13) имеет смысл, т. е. $g(a) \neq g(b)$. В самом деле, если $g(a) = g(b)$, то функция g удовлетворяла бы условиям теоремы Ролля и, значит, нашлась бы такая точка ξ, что $g'(\xi) = 0$, $a < \xi < b$, что противоречило бы условию 3.

Доказательство. Рассмотрим вспомогательную функцию

$$F(x) = f(x) - \lambda g(x), \qquad (11.14)$$

где число λ выберем таким образом, чтобы $F(a) = F(b)$, т. е. чтобы $f(a) - \lambda g(a) = f(b) - \lambda g(b)$. Для этого нужно взять

$$\lambda = \frac{f(b) - f(a)}{g(b) - g(a)}. \qquad (11.15)$$

Функция F удовлетворяет всем условиям теоремы Ролля, следовательно, существует такая точка ξ, $a < \xi < b$, что $F'(\xi) = 0$. Но из (11.14) $F'(x) = f'(x) - \lambda g'(x)$, поэтому

$$f'(\xi) - \lambda g'(\xi) = 0,$$

откуда следует, что

$$\lambda = \frac{f'(\xi)}{g'(\xi)}. \qquad (11.16)$$

Сравнив (11.15) и (11.16), получим формулу (11.13), обычно называемую *формулой конечных приращений Коши.* □

Отметим, что формула конечных приращений Лагранжа является частным случаем формулы конечных приращений Коши, в которой $g(x) = x$. Мы привели независимые доказательства этих формул, во-первых, из-за той важной роли, которую играет формула Лагранжа, а во-вторых, чтобы иметь возможность, используя одну и ту же идею (построения вспомогательной функции, удовлетворяющей условиям теоремы Ролля), применить ее дважды в доказательствах, причем сначала, для большей наглядности, в более простом случае.

Формула Коши (11.13), так же как и формула Лагранжа (11.3), справедлива не только если $a < b$, но и если $a > b$.

Выше было показано (см. п. 6.2), что всякая непрерывная на отрезке функция, принимая какие-либо два значения, принимает и любое промежуточное значение. Оказыва-

ется, что тем же свойством обладают производные функций, хотя они могут быть и разрывными функциями.

Функция, у которой производная разрывна, приведена в примере 9 п. 9.7:

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x} & \text{при } x \neq 0, \\ 0 & \text{при } x = 0. \end{cases}$$

Она имеет производную во всех точках числовой оси:

$$f'(x) = \begin{cases} 2x \sin \frac{1}{x} - \cos \frac{1}{x} & \text{при } x \neq 0, \\ 0 & \text{при } x = 0, \end{cases}$$

разрывную в точке $x = 0$.

ТЕОРЕМА 5 (Дарбу[1]). *Если функция дифференцируема в каждой точке отрезка, то производная этой функции, принимая какие-либо два значения, принимает и любое промежуточное.*

Доказательство. Рассмотрим сначала случай, когда функция f имеет в каждой точке отрезка $[a, b]$ производную, принимающую на концах отрезка значения разного знака. Здесь, как всегда, раз не оговорено что-либо другое, под производной понимается конечная производная, а на концах отрезка рассматриваются односторонние производные.

Пусть, например,

$$f'(a) > 0, \ f'(b) < 0.$$

Из условия

$$f'(a) = \lim_{h \to +0} \frac{f(a+h) - f(a)}{h} > 0$$

следует, что существует такое $\delta > 0$, что для всех h, $0 < h < \delta$, выполняется условие $\frac{f(a+h) - f(a)}{h} > 0$, а так как $h > 0$, то $f(a) < f(a + h)$.

Аналогично из условия

$$f'(b) = \lim_{h \to -0} \frac{f(b+h) - f(b)}{h} < 0$$

следует, что в некоторой окрестности точки $x = b$ выполняется неравенство $\frac{f(b+h) - f(b)}{h} < 0$, где $h < 0$, и, следовательно, $f(b) < f(b + h)$.

---

[1] Г. Дарбу (1842—1917) — французский математик.

Таким образом, функция f заведомо не принимает своего наибольшего значения в точках $x = a$ и $x = b$. Однако функция f, будучи дифференцируемой на отрезке $[a, b]$, является и непрерывной на нем, поэтому согласно теореме Вейерштрасса (см. теорему 1 в п. 6.1) принимает в некоторой точке $\xi \in [a, b]$ свое наибольшее значение. Согласно доказанному $\xi \neq a$ и $\xi \neq b$, следовательно, $\xi \in (a, b)$. Тогда, согласно теореме Ферма (см. теорему 1 в п. 11.1), имеем $f'(\xi) = 0$.

Аналогично рассматривается случай, когда $f'(a) < 0$, $f'(b) > 0$.

Если теперь функция f дифференцируема на отрезке $[a, b]$, $f'(a) < c < f'(b)$, то функция $F(x) = f(x) - cx$ также дифференцируема на этом отрезке, $F'(x) = f'(x) - c$ и, следовательно, $F'(a) = f'(a) - c < 0$, $F'(b) = f'(b) - c > 0$. Поэтому по доказанному выше существует такая точка $\xi \in (a, b)$, что $F'(\xi) = 0$, что равносильно тому, что $f'(\xi) = c$.

Аналогично рассматривается случай $f'(a) > c > f'(b)$. □

УПРАЖНЕНИЯ. 4. Пусть $f(x) = x^2 \sin \frac{1}{x}$ при $x \neq 0$, $f(0) = 0$. Применим к этой функции на отрезке $[0, x]$ формулу Лагранжа:

$$x^2 \sin \frac{1}{x} = \left(2\xi \sin \frac{1}{\xi} - \cos \frac{1}{\xi}\right)x,$$

где $0 < \xi < x$. Сократим обе части равенства на x при $x \neq 0$:

$$x \sin \frac{1}{x} = 2\xi \sin \frac{1}{\xi} - \cos \frac{1}{\xi}.$$

Переходя к пределу при $x \to 0$ (при этом, очевидно, $\xi \to 0$), получаем

$$\lim_{\xi \to 0} \cos \frac{1}{\xi} = 0,$$

так как два других слагаемых, очевидно, стремятся к нулю. Вместе с тем предел функции $\cos \frac{1}{\xi}$ при стремлении аргумента к нулю не существует! Где ошибка?

5. Доказать, что если функция дифференцируема на интервале и ее производная принимает значения разных знаков, то существует точка, в которой производная обращается в нуль.

6. Доказать, что если функция дифференцируема на интервале и ее производная не обращается на нем в нуль, то функция строго монотонна.

Задача 8. Доказать, что многочлен

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n(x^2 - 1)^n}{dx^n}$$

(называемый многочленом Лежандра[1]) имеет на интервале $(-1, 1)$ n различных действительных корней.

Задача 9. Доказать, что многочлен

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n e^{-x^2}}{dx^n}$$

(называемый многочленом Чебышева—Эрмита[2]) имеет n различных действительных корней, лежащих на интервале $(-\sqrt{2n+1}, \sqrt{2n+1})$, $n = 1, 2, \dots$.

§ 12.

Раскрытие неопределенностей по правилу Лопиталя

Во многих случаях отыскание предела функции, заданной аналитически, при стремлении аргумента к некоторой точке (числу или к одной из бесконечностей ∞, $+\infty$ или $-\infty$), выполняемое формальной подстановкой соответствующего значения вместо аргумента в формулу, задающую рассматриваемую функцию, приводит к выражениям вида $\frac{0}{0}$, $\frac{\infty}{\infty}$, $0 \cdot \infty$, $\infty - \infty$, 0^0, ∞^0 или 1^∞. Они называются *неопределенностями*, так как по ним нельзя судить о том, существует или нет указанный предел, не говоря уже о нахождении его значения (если он существует). В этом случае вычисление предела называется также *раскрытием неопределенности*.

Наряду с основным приемом нахождения пределов функции — методом выделения главной части — существуют и другие способы отыскания пределов. Некоторые из них, носящие общее название *правил Лопиталя*[3], мы изложим в этом параграфе.

---

[1] А. М. Лежандр (1752—1833) — французский математик.

[2] П. Л. Чебышев (1821—1894) — русский математик и механик; Ш. Эрмит (1822—1894) — французский математик.

[3] Г. Лопиталь (1661—1704) — французский математик.

12.1. Неопределенности вида $\frac{0}{0}$

ТЕОРЕМА 1. *Пусть функции* $f(x)$ *и* $g(x)$*, определенные на отрезке* $[a, b]$*, таковы, что*:
1) $f(a) = g(a) = 0$;
2) *существуют производные (правосторонние)* $f'(a)$ *и* $g'(a)$*, причем* $g'(a) \neq 0$.

Тогда существует предел

$$\lim_{x \to a+0} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)}.$$

Доказательство. Применим метод выделения главной части. В силу условия 2, имеем (см. п. 9.2)

$$f(x) = f(a) + f'(a)(x - a) + o(x - a),$$
$$g(x) = g(a) + g'(a)(x - a) + o(x - a).$$

Отсюда, согласно условию 1, получим, что

$$f(x) = f'(a)(x - a) + o(x - a),$$
$$g(x) = g'(a)(x - a) + o(x - a),$$

поэтому

$$\lim_{x \to a+0} \frac{f(x)}{g(x)} = \lim_{x \to a+0} \frac{f'(a) + \frac{o(x-a)}{x-a}}{g'(a) + \frac{o(x-a)}{x-a}} = \frac{f'(a)}{g'(a)}. \square$$

В теореме 1 предполагалось существование производных в точке a. Докажем теперь теорему, близкую по содержанию к предыдущей, в которой, однако, не будет предполагаться существование производных $f'(a)$ и $g'(a)$.

ТЕОРЕМА 2. *Пусть функции* $f(x)$ *и* $g(x)$:
1) *дифференцируемы на интервале* (a, b);
2) $\lim\limits_{x \to a+0} f(x) = \lim\limits_{x \to a+0} g(x) = 0$;
3) $g'(x) \neq 0$ *для всех* $x \in (a, b)$;
4) *существует конечный или бесконечный, равный* $+\infty$ *или* $-\infty$*, предел* $\lim\limits_{x \to a+0} \frac{f'(x)}{g'(x)}$.

Тогда существует предел $\lim\limits_{x \to a+0} \frac{f(x)}{g(x)} = \lim\limits_{x \to a+0} \frac{f'(x)}{g'(x)}$.

Доказательство. В силу условий теоремы, функции f и g не определены в точке a; доопределим их, положив $f(a) = g(a) = 0$. Теперь f и g непрерывны в точке a и удовлетворя-

ют условиям теоремы Коши о среднем значении (см. п. 11.2) на любом отрезке $[a, x]$, где $a < x < b$. Поэтому для каждого x, $a < x < b$, существует такое $\xi = \xi(x) \in (a, x)$, что

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\xi)}{g'(\xi)}, \tag{12.1}$$

причем $\lim\limits_{x \to a+0} \xi(x) = a$.

Поэтому, если существует $\lim\limits_{x \to a+0} \frac{f'(x)}{g'(x)} = k$, то из правила замены переменного для пределов функций следует, что существует и $\lim\limits_{x \to a+0} \frac{f'(\xi)}{g'(\xi)} = k$. Теперь из (12.1) получаем

$$\lim_{x \to a+0} \frac{f(x)}{g(x)} = \lim_{x \to a+0} \frac{f'(\xi)}{g'(\xi)} = k. \ \Box$$

Теоремы 1 и 2 остаются верными с естественными видоизменениями как в случае левостороннего, так и двустороннего предела.

ТЕОРЕМА 3. *Пусть функции f и g:*

1) *дифференцируемы при $x > c$;*

2) $\lim\limits_{x \to +\infty} f(x) = 0$, $\lim\limits_{x \to +\infty} g(x) = 0$;

3) $g'(x) \neq 0$ *для всех $x > c$;*

4) *существует конечный или бесконечный, равный $+\infty$ или $-\infty$, предел* $\lim\limits_{x \to +\infty} \frac{f'(x)}{g'(x)}$.

Тогда существует и предел $\lim\limits_{x \to +\infty} \frac{f(x)}{g(x)} = \lim\limits_{x \to +\infty} \frac{f'(x)}{g'(x)}$.

Доказательство. Без ограничения общности можно считать, что $c > 0$ (если $c < 0$, то в качестве нового значения c возьмем, например, $c = 1$).

Выполним замену переменного $x = 1/t$. Функции $\varphi(t) = f(1/t)$ и $\psi(t) = g(1/t)$ определены на интервале $(0, 1/c)$; если $x \to +\infty$, то $t \to +0$ и наоборот. На интервале $(0, 1/c)$ существуют производные

$$\varphi'(t) = -f'\left(\frac{1}{t}\right)\frac{1}{t^2} \ \text{ и } \ \psi'(t) = -g'\left(\frac{1}{t}\right)\frac{1}{t^2},$$

где штрихом обозначены производные функций f и g по первоначальному аргументу.

Из сказанного и условий теоремы следует, что функции $\varphi(t)$ и $\psi(t)$ удовлетворяют на интервале $(0, 1/c)$ условиям 1), 2) и 3) теоремы 2. Покажем еще, что из существования предела $\lim\limits_{x \to +\infty} \frac{f'(x)}{g'(x)}$, который обозначим через k, следует существование предела $\lim\limits_{x \to +0} \frac{\varphi'(t)}{\psi'(t)}$ и равенство его k, т. е. что выполняется и условие 4) теоремы 2. Действительно, используя полученные выражения для производных $\varphi'(t)$ и $\psi'(t)$, находим

$$\lim_{t \to +0} \frac{\varphi'(t)}{\psi'(t)} = \lim_{t \to +0} \frac{f'(1/t)}{g'(1/t)} = \lim_{x \to +\infty} \frac{f'(x)}{g'(x)} = k.$$

Теперь из теоремы 2, примененной к функциям $\varphi(t)$ и $\psi(t)$, следует, что $\lim\limits_{t \to +0} \frac{\varphi(t)}{\psi(t)} = k$. Но

$$\frac{\varphi(t)}{\psi(t)} = \frac{f(1/t)}{g(1/t)} = \frac{f(x)}{g(x)},$$

где $x = 1/t$, поэтому

$$\lim_{t \to +\infty} \frac{f(x)}{g(x)} = \lim_{t \to +0} \frac{\varphi'(t)}{\psi'(t)} = k. \ \square$$

Эта теорема остается верной с соответствующим видоизменением и при $x \to -\infty$.

Правилом Лопиталя называется нахождение предела отношений функций по формуле

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)},$$

где a — конечная или бесконечно удаленная точка числовой оси, а функции f и g заданы во всех точках некоторой ее двусторонней или односторонней окрестности, кроме самой этой точки.

В случае конечной точки a правилом Лопиталя называется также и нахождение предела отношений функций по формуле

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)}.$$

12.2. Неопределенности вида $\frac{\infty}{\infty}$

ТЕОРЕМА 4. *Пусть функции* $f(x)$ *и* $g(x)$:

1) *дифференцируемы на интервале* (a, b);

2) $\lim\limits_{x \to a+0} f(x) = \infty$, $\lim\limits_{x \to a+0} g(x) = \infty$;

3) $g'(x) \neq 0$ *на* (a, b);

4) *существует конечный или бесконечный, равный* $+\infty$ *или* $-\infty$, *предел* $\lim\limits_{x \to a+0} \dfrac{f'(x)}{g'(x)}$.

Тогда существует и предел

$$\lim_{x \to a+0} \frac{f(x)}{g(x)} = \lim_{x \to a+0} \frac{f'(x)}{g'(x)}.$$

Доказательство. Пусть существует конечный или бесконечный предел

$$\lim_{x \to a+0} \frac{f'(x)}{g'(x)} = k. \tag{12.3}$$

Покажем, что при выполнении остальных условий теоремы имеет место равенство

$$\lim_{x \to a+0} \frac{f(x)}{g(x)} = k. \tag{12.4}$$

Если $a < x < x_0 < b$, то на отрезке $[x, x_0]$ функции f и g удовлетворяют условиям теоремы Коши (см. теорему 4 в п. 11.2), поэтому существует такая точка $\xi = \xi(x_0, x)$, что

$$\frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{f'(\xi)}{g'(\xi)}, \quad x < \xi < x_0. \tag{12.5}$$

Далее, в силу (12.2), существует такая точка $x_1 = x_1(x_0)$, $a < x_1 < x_0$, что при всех $x \in (a, x_1)$ выполняются неравенства

$$f(x) \neq 0,\ g(x) \neq 0,\ f(x) \neq f(x_0) \tag{12.6}$$

и, следовательно, можно выполнять деление на $f(x)$, $g(x)$ и $1 - \dfrac{f(x_0)}{f(x)}$ (а также и на $1 - \dfrac{g(x_0)}{g(x)}$, поскольку, в силу условий теоремы, $g(x) \neq g(x_0)$, см. теорему 4 в п. 11.2). Для этих значений x из (12.5) вытекает равенство

$$\frac{f(x)}{g(x)} \frac{1 - \dfrac{f(x_0)}{f(x)}}{1 - \dfrac{g(x_0)}{g(x)}} = \frac{f'(\xi)}{g'(\xi)},$$

откуда

$$\frac{f(x)}{g(x)} = \frac{f'(\xi)}{g'(\xi)} \frac{1 - \dfrac{g(x_0)}{g(x)}}{1 - \dfrac{f(x_0)}{f(x)}}. \tag{12.7}$$

В правой части равенства первый сомножитель $\frac{f'(\xi)}{g'(\xi)}$ стремится к k при $x_0 \to a + 0$, так как $a < \xi < x_0$, а второй, в силу условия (12.2), стремится к 1 при $x \to a + 0$ и фиксированном x_0:

$$\lim_{x \to a + 0} \frac{1 - \frac{g(x_0)}{g(x)}}{1 - \frac{f(x_0)}{f(x)}} = 1. \qquad (12.8)$$

Непосредственно перейти к пределу в равенстве (12.7) нельзя, поскольку указанные выше предельные переходы в сомножителях в правой части равенства происходят при разных условиях: при $x_0 \to a + 0$ и при фиксированном x_0, но $x \to a + 0$. Однако если задать произвольно окрестность $U(k)$ предела k, то, в силу условия (12.3), можно сначала зафиксировать точку x_0 столь близко к точке a, что отношение $\frac{f'(\xi)}{g'(\xi)}$ попадет в эту окрестность, ибо $a < \xi < x_0$. Согласно же условию (12.8), для всех точек x, достаточно близких к a, отношение $\frac{f(x)}{g(x)}$ (см. (12.7)) также будет принадлежать указанной окрестности $U(k)$, а это означает справедливость утверждения (12.4). Собственно говоря, теорема доказана и можно поставить знак □.

Однако для полноты изложения сделаем некоторые дополнительные разъяснения, которые каждый, кто достаточно хорошо овладел предшествующим материалом, легко может провести самостоятельно.

Пусть сначала k — конечный предел. Зададим произвольно $\varepsilon > 0$ и так зафиксируем x_0, чтобы кроме условий (12.6) для всех x, $a < x < x_0$, выполнялось неравенство

$$\left|\frac{f'(x)}{g'(x)} - k\right| < \frac{\varepsilon}{3}. \qquad (12.9)$$

(Это возможно в силу условия (12.3).) Таким образом, если положить

$$\alpha(x) \overset{\text{def}}{=} \frac{f'(x)}{g'(x)} - k, \qquad (12.10)$$

то при $a < x < x_0$ выполняется неравенство

$$|\alpha(x)| < \frac{\varepsilon}{3}. \qquad (12.11)$$

Если

$$\beta(x) \stackrel{\text{def}}{=} \frac{1 - \dfrac{g(x_0)}{g(x)}}{1 - \dfrac{f(x_0)}{f(x)}} - 1, \tag{12.12}$$

то, согласно (12.8),

$$\lim_{x \to a+0} \beta(x) = 0, \tag{12.13}$$

поэтому существует такое x_1, $a < x_1 < x_0$, что для всех точек $x \in (a, x_1)$ имеют место неравенства

$$|\beta(x)| < 1, \ |\beta(x)| < \frac{1}{3(|k|+1)}\varepsilon. \tag{12.14}$$

В силу формул (12.7), (12.10), (12.12) имеем

$$\frac{f(x)}{g(x)} = (k + \alpha(\xi))(1 + \beta(x)),$$

поэтому

$$\frac{f(x)}{g(x)} - k = \alpha(\xi) + k\beta(x) + \alpha(\xi)\beta(x), \quad a < \xi < x_0,$$

где при $a < x < x_1$ имеем

$$|\alpha(\xi) + k\beta(x) + \alpha(\xi)\beta(x)| \leqslant$$
$$\leqslant |\alpha(\xi)| + |k||\beta(x)| + |\alpha(\xi)||\beta(x)| \underset{\substack{(12.11)\\(12.14)}}{<} \frac{\varepsilon}{3} + \frac{|k|}{3(|k|+1)}\varepsilon + \frac{\varepsilon}{3} < \varepsilon,$$

а это и означает справедливость равенства (12.4).

Так раскрываются на «языке неравенств» сделанные выше высказывания о выборе достаточно близких значений x_0 и x к точке a, обеспечивающих нужную близость отношения $\frac{f(x)}{g(x)}$ к числу k.

Рассмотрим теперь случай бесконечного предела. Пусть $\lim\limits_{x \to a+0} \frac{f'(x)}{g'(x)} = +\infty$. Тогда в некоторой окрестности точка a имеем $f'(x) \neq 0$ (почему?) и, следовательно, $\lim\limits_{x \to a+0} \frac{g'(x)}{f'(x)} = 0$. Поэтому, согласно доказанному выше, $\lim\limits_{x \to a+0} \frac{g(x)}{f(x)} = 0$, откуда следует, что

$$\lim_{x \to a+0} \frac{f(x)}{g(x)} = \infty.$$

Но нужно доказать более сильное утверждение, а именно что этот предел равен $+\infty$. Покажем это.

Зададим произвольно число $c > 0$. Из условия (12.3) при $k = +\infty$ следует, что существует такое $x_0 \in (a, b)$, для которого при всех $x \in (a, x_0)$, кроме условий (12.6), будет выполняться еще неравенство

$$\frac{f'(x)}{g'(x)} > 2c. \tag{12.15}$$

В силу же условия (12.8) или равносильного ему условия (12.13), существует такое x_1, $a < x_1 < x_0$, что для всех x, $a < x < x_1$, выполняется неравенство

$$1 + \beta(x) > \frac{1}{2}. \tag{12.16}$$

Теперь для всех $x \in (a, x_1)$ имеем (вспомним, что $a < \xi < x_0$)

$$\frac{f(x)}{g(x)} \underset{(12.7)}{=} \frac{f'(\xi)}{g'(\xi)} \frac{1 - \dfrac{g(x_0)}{g(x)}}{1 - \dfrac{f(x_0)}{f(x)}} \underset{(12.12)}{=} \frac{f'(\xi)}{g'(\xi)}(1 + \beta(x)) \underset{\substack{(12.15)\\(12.16)}}{>} 2c \cdot \frac{1}{2} = c.$$

Это означает, что $\lim\limits_{x \to a+0} \frac{f(x)}{g(x)} = +\infty$. Случай $k = -\infty$ сводится к случаю $k = +\infty$ заменой функции $f(x)$ на функцию $-f(x)$.

Теорема 4 вместе с ее доказательством остается в силе с естественными видоизменениями и при $x \to a - 0$, $x \to +\infty$ и $x \to -\infty$, а также в случае двусторонних пределов.

Можно показать, что при выполнении условий 1), 2) и 3), входящих в любую из теорем 2, 3 или 4, не может существовать предел $\lim\limits_{x \to a+0} \frac{f'(x)}{g'(x)} = \infty$ без существования одного из двух «знакоопределенных» бесконечных пределов $\lim\limits_{x \to a+0} \frac{f'(x)}{g'(x)} = +\infty$ или $\lim\limits_{x \to a+0} \frac{f'(x)}{g'(x)} = -\infty$ (см. далее упражнение 2 в п. 12.3). Таким образом, рассмотрены все случаи неопределенностей вида $\frac{\infty}{\infty}$.

Примеры. 1. Найдем $\lim\limits_{x \to +\infty} \frac{\ln x}{x^\alpha}$, $\alpha > 0$. Замечая, что

$$(\ln x)' = \frac{1}{x}, \ (x^\alpha)' = \alpha x^{\alpha - 1} \text{ и } \lim_{x \to +\infty} \frac{1/x}{\alpha x^{\alpha - 1}} = \lim_{x \to +\infty} \frac{1}{x^\alpha} = 0,$$

получаем

$$\lim_{x \to +\infty} \frac{\ln x}{x^{\alpha}} = 0.$$

Это означает, что при $x \to +\infty$ функция $\ln x$ растет медленнее, чем любая положительная степень переменной x.

Иногда правило Лопиталя приходится применять несколько раз.

2. Найдем $\lim\limits_{x \to +\infty} \frac{x^n}{a^x}$, где n — натуральное число и $a > 1$.

Имеем

$$\lim_{x \to +\infty} \frac{x^n}{a^x} = \lim_{x \to +\infty} \frac{(x^n)'}{(a^x)'} = \lim_{x \to +\infty} \frac{nx^{n-1}}{a^x \ln a} = \ldots =$$

$$= \lim_{x \to +\infty} \frac{n!}{a^x \ln^n a} = 0. \qquad (12.17)$$

Таким образом, при $x \to +\infty$ любая степень x^n растет медленнее, чем показательная функция a^x, $a > 1$.

3. Следует иметь в виду, что вычисления по правилу Лопиталя оправданы только в том случае, когда в результате получается конечный или бесконечный предел. Так, например, было бы ошибкой написать

$$\lim_{x \to \infty} \frac{x - \sin x}{x + \sin x} = \lim_{x \to \infty} \frac{(x - \sin x)'}{(x + \sin x)'},$$

так как предел

$$\lim_{x \to \infty} \frac{(x - \sin x)'}{(x + \sin x)'} = \lim_{x \to \infty} \frac{1 - \cos x}{1 + \cos x}$$

не существует.

В самом деле, беря последовательности $x'_n = 2\pi n \to +\infty$ и $x''_n = \frac{\pi}{2} + 2\pi n \to +\infty$ при $n \to \infty$, получаем

$$\lim_{n \to \infty} \frac{1 - \cos x'_n}{1 + \cos x'_n} = 0, \quad \lim_{n \to \infty} \frac{1 - \cos x''_n}{1 + \cos x''_n} = 1.$$

Вместе с тем заданная в этом случае неопределенность $\frac{\infty}{\infty}$ может быть раскрыта элементарно:

$$\lim_{x \to \infty} \frac{x - \sin x}{x + \sin x} = \lim_{x \to \infty} \frac{1 - \dfrac{\sin x}{x}}{1 + \dfrac{\sin x}{x}} = 1.$$

УПРАЖНЕНИЕ 1. Пусть $f(x) = x^2 \sin \frac{1}{x}$, $g(x) = \sin x$. Найти $\lim\limits_{x \to 0} \frac{f(x)}{g(x)}$ и доказать, что в этом случае правило Лопиталя неприменимо.

4. Может случиться, что применение правила Лопиталя не упрощает задачу отыскания предела функции. Например, применив правило Лопиталя для вычисления предела $\lim\limits_{x \to +\infty} \frac{x}{\sqrt{1+x^2}}$, получим

$$\lim_{x \to +\infty} \frac{x}{\sqrt{1+x^2}} = \lim_{x \to +\infty} \frac{x'}{(\sqrt{1+x^2})'} = \lim_{x \to +\infty} \frac{\sqrt{1+x^2}}{x},$$

т. е. получился предел дроби, обратной данной, т. е. задача осталась той же. Вместе с тем заданный предел легко находится элементарно:

$$\lim_{x \to +\infty} \frac{x}{\sqrt{1+x^2}} = \lim_{x \to +\infty} \frac{1}{\sqrt{1+1/x^2}} = 1.$$

5. Неопределенности 0^0, ∞^0 или 1^∞ можно раскрыть, предварительно прологарифмировав соответствующие функции. Например, чтобы найти $\lim\limits_{x \to +0} x^x$, следует найти предел:

$$\lim_{x \to +0} x \ln x = \lim_{x \to +0} \frac{\ln x}{1/x} = -\lim \frac{1/x}{1/x^2} = -\lim_{x \to +0} x = 0.$$

Поэтому, в силу непрерывности показательной функции,

$$\lim_{x \to +0} x^x = \lim_{x \to +0} e^{x \ln x} = 1.$$

Неопределенности вида $0 \cdot \infty$ и $\infty - \infty$ следует привести к виду $\frac{0}{0}$ или $\frac{\infty}{\infty}$. При этом, как и всегда при применении правила Лопиталя, по ходу вычислений рекомендуется упрощать получающиеся выражения. Поясним это на примере.

6. $\lim\limits_{x \to 0} \left(\frac{1}{x^2} - \operatorname{ctg}^2 x\right) = \lim\limits_{x \to 0} \frac{\sin^2 x - x^2 \cos^2 x}{x^2 \sin^2 x}$.

Заметим, что

$$\frac{\sin^2 x - x^2 \cos^2 x}{x^2 \sin^2 x} = \frac{\sin x + x \cos x}{\sin x} \cdot \frac{\sin x - x \cos x}{x^2 \sin x}.$$

Предел первого сомножителя правой части находится непосредственно:

$$\lim_{x \to 0} \frac{\sin x + x \cos x}{\sin x} = \lim_{x \to 0} \left(1 + \frac{x}{\sin x} \cos x\right) = 2,$$

а предел второго — с помощью правил Лопиталя:

$$\lim_{x \to 0} \frac{\sin x - x\cos x}{x^2 \sin x} = \lim_{x \to 0} \frac{x \sin x}{2x \sin x + x^2 \cos x} =$$
$$= \lim_{x \to 0} \frac{1}{2 + \frac{x}{\sin x} \cos x} = \frac{1}{3}.$$

Таким образом, $\lim\limits_{x \to 0} \left(\frac{1}{x^2} - \operatorname{ctg}^2 x \right) = \frac{2}{3}$.

12.3. Обобщение правила Лопиталя

Усовершенствовав доказательство теоремы 4, можно показать справедливость правила Лопиталя при более слабых ограничениях. Рассмотрим для разнообразия случай, когда аргумент стремится к $+\infty$. Приведенное ниже изложение доказательства теоремы 5 принадлежит В. А. Ходакову.

ТЕОРЕМА 5. *Пусть функции* $f(x)$ *и* $g(x)$:

1) *дифференцируемы при* $x > a$;

2) $g'(x) \neq 0$ *при* $x > a$; (12.18)

3) $\lim\limits_{x \to +\infty} g(x) = \infty$; (12.19)

4) *существует конечный или определенного знака бесконечный предел* $\lim\limits_{x \to +\infty} \frac{f'(x)}{g'(x)} = k$.

Тогда

$$\lim_{x \to +\infty} \frac{f(x)}{g(x)} = k. \tag{12.20}$$

Как и выше, можно показать, что из условий (12.18)—(12.20) следует, что если $\lim\limits_{x \to +\infty} \frac{f'(x)}{g'(x)} = \infty$, то либо $\lim\limits_{x \to +\infty} \frac{f'(x)}{g'(x)} = +\infty$, либо $\lim\limits_{x \to +\infty} \frac{f'(x)}{g'(x)} = -\infty$. Поэтому случай $k = \infty$ рассматривать не надо.

Доказательство. 1) Пусть сначала $k \neq \pm\infty$. Зафиксируем произвольно $\varepsilon > 0$. Выберем такое $x_0 > a$, чтобы для всех $x > x_0$ выполнялись неравенства $g(x) \underset{(12.19)}{\neq} 0$ и

$$\left| \frac{f'(x)}{g'(x)} - k \right| < \frac{\varepsilon}{2}. \tag{12.21}$$

Согласно теореме Коши, для каждого $x > x_0$ существует такое $\xi = \xi(x, x_0) > x_0$, что

$$f(x) = \frac{f'(\xi)}{g'(\xi)}[g(x) - g(x_0)] + f(x_0).$$

Отсюда

$$\frac{f(x)}{g(x)} - k = \frac{f'(\xi)}{g'(\xi)} - k + \frac{1}{g(x)}\left[f(x_0) - \frac{f'(\xi)}{g'(\xi)}g(x_0)\right].$$

Положив

$$\alpha(x) \stackrel{\text{def}}{=} \frac{1}{|g(x)|}\left|f(x_0) - \frac{f'(\xi)}{g'(\xi)}g(x_0)\right|,$$

будем иметь

$$\left|\frac{f(x)}{g(x)} - k\right| \leqslant \left|\frac{f'(\xi)}{g'(\xi)} - k\right| + \alpha(x). \qquad (12.22)$$

В силу условия (12.21), отношение $\frac{f'(\xi)}{g'(\xi)}$ ограничено при $x > x_0$. Следовательно, $\lim\limits_{x \to +\infty} \alpha(x) \underset{(12.19)}{=} 0$. Поэтому найдется такое $x_1 > x_0$, что для всех $x > x_1$ будет выполняться неравенство

$$\alpha(x) < \frac{\varepsilon}{2} \qquad (12.23)$$

и, таким образом, при $x > x_1$ верно неравенство

$$\left|\frac{f(x)}{g(x)} - k\right| \underset{(12.22)}{\leqslant} \left|\frac{f'(\xi)}{g'(\xi)} - k\right| + \alpha(x) \underset{\substack{(12.21)\\(12.23)}}{<} \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

а это и означает, что $\lim\limits_{x \to +\infty} \frac{f(x)}{g(x)} = k$.

2) Если, например,

$$\lim_{x \to +\infty} \frac{f'(x)}{g'(x)} = +\infty, \qquad (12.24)$$

то для любого $c > 0$ существует такое $x_0 > a$, что для всех $x > x_0$ выполняются неравенства $g(x) \neq 0$ и

$$\frac{f'(x)}{g'(x)} > 3c. \qquad (12.25)$$

Согласно теореме Коши, при любом $x > x_0$ имеет место равенство

$$\frac{f(x)}{g(x)} = \frac{f'(\xi)}{g'(\xi)}\left[1 - \frac{g(x_0)}{g(x)}\right] + \frac{f(x_0)}{g(x)},$$

где $\xi = \xi(x_0, x) > x_0$. В силу условия (12.19), существует такое $x_1 > x_0$, что для всех $x > x_1$ выполняются неравенства

$$1 - \frac{g(x_0)}{g(x)} > \frac{2}{3}, \quad \left|\frac{f(x_0)}{g(x)}\right| < c, \tag{12.26}$$

следовательно, и неравенство

$$\frac{f(x)}{g(x)} \underset{\substack{(12.25)\\(12.26)}}{>} 3c \cdot \frac{2}{3} - c = c.$$

Это и означает, что $\lim\limits_{x \to +\infty} \frac{f(x)}{g(x)} = +\infty$.

Утверждение теоремы для $k = -\infty$ следует из случая $k = +\infty$, надо лишь $f(x)$ заменить на $-f(x)$. □

УПРАЖНЕНИЕ 2. Доказать, что если функции $f(x)$ и $g(x)$ дифференцируемы на интервале (a, b), $g'(x) \neq 0$ на (a, b), и $\lim\limits_{x \to +\infty} \frac{f'(x)}{g'(x)} = \infty$, то либо $\lim\limits_{x \to +\infty} \frac{f'(x)}{g'(x)} = +\infty$, либо $\lim\limits_{x \to +\infty} \frac{f'(x)}{g'(x)} = -\infty$.

Указание. Полезно воспользоваться упражнениями 5 и 6 в п. 11.2.

§ 13.

Формула Тейлора

13.1. Вывод формулы Тейлора

Если функция $y = f(x)$ имеет в точке x_0 производную, то приращение этой функции можно представить в виде

$$\Delta y = A\,\Delta x + o\,(\Delta x),\ \ \Delta x \to 0,$$

где $\Delta x = x - x_0$, $\Delta y = f(x) - y_0$, $y_0 = f(x_0)$ и $A = f'(x_0)$, т. е.

$$f(x) = y_0 + A(x - x_0) + o(x - x_0),\ x \to x_0.$$

Иначе говоря, существует линейная функция

$$P_1(x) = y_0 + A(x - x_0) \tag{13.1}$$

такая, что

$$f(x) = P_1(x) + o(x - x_0),\ x \to x_0,$$

причем

$$P_1(x_0) = y_0 = f(x_0), \quad P_1'(x_0) = A = f'(x_0).$$

Поставим более общую задачу. Пусть функция f имеет n производных в точке x_0. Требуется выяснить, существует ли многочлен $P_n(x)$ степени не выше n такой, что

$$f(x) = P_n(x) + o((x - x_0)^n),\ x \to x_0, \tag{13.2}$$

и

$$f(x_0) = P_n(x_0),\ f'(x_0) = P_n'(x_0),\ \ldots,\ f^{(n)}(x_0) = P_n^{(n)}(x_0). \tag{13.3}$$

Будем искать этот многочлен, по аналогии с формулой (13.1), в виде

$$P_n(x) = A_0 + A_1(x - x_0) + A_2(x - x_0)^2 + \ldots + A_n(x - x_0)^n.$$

Замечая, что $P_n(x_0) = A_0$ из первого условия (13.3), т. е. условия $f(x_0) = P_n(x_0)$, имеем $A_0 = f(x_0)$. Далее,

$$P_n'(x) = A_1 + 2A_2(x - x_0) + \ldots + nA_n(x - x_0)^{n-1},$$

отсюда $P_n'(x_0) = A_1$, и так как $P_n'(x_0) \underset{(13.3)}{=} f'(x_0)$, то $A_1 = f'(x_0)$. Затем найдем вторую производную многочлена $P_n(x)$:

$$P_n''(x) = 2 \cdot 1 \cdot A_2 + \ldots + n(n-1)A_n(x - x_0)^{n-2}.$$

Отсюда и из условия $f''(x_0) \underset{(13.3)}{=} P_n''(x_0)$ получим $A_2 = \frac{f''(x_0)}{2!}$ и вообще

$$A_k = \frac{f^{(k)}(x_0)}{k!},\ k = 0, 1, 2, \ldots, n.$$

В силу самого построения, для многочлена

$$P_n(x) = f(x_0) + f'(x_0)(x - x_0) + \ldots$$
$$\ldots + \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

выполнены все соотношения (13.3). Проверим, удовлетворяет ли он условию (13.2).

Пусть

$$r_n(x) \overset{\text{def}}{=} f(x) - P_n(x).$$

Как уже отмечалось выше (см. п. 10.1), из существования у функции f производной порядка n в точке x_0 следует, согласно определению этой производной, что у функции f в некоторой окрестности точки x_0 существует производная порядка $n-1$, следовательно, и все производные более низкого порядка. При этом, поскольку функция, имеющая в некоторой точке производную, непрерывна в этой точке, все производные функции f до порядка $n-1$ включительно непрерывны в точке x_0. Поэтому у функции $r_n(x)$ также в точке x_0 существуют производные до порядка n, а в некоторой окрестности этой точки все производные до порядка $n-1$ включительно и они непрерывны в точке x_0.

Из условия (13.3) следует, что

$$r_n(x_0) = r'_n(x_0) = \ldots = r_n^{(n)}(x_0) = 0. \tag{13.4}$$

В силу свойств производных функции $r_n(x)$, для раскрытия неопределенности $\frac{r_n(x)}{(x-x_0)^n}$ при $x \to x_0$ можно применить правило Лопиталя, а именно, сначала $n-1$ раз теорему 2 из § 12, а затем теорему 1 из того же параграфа. В результате получим

$$\lim_{x \to x_0} \frac{r_n(x_0)}{(x-x_0)^n} = \lim_{x \to x_0} \frac{r'_n(x)}{n(x-x_0)^{n-1}} = \ldots$$

$$\ldots = \lim_{x \to x_0} \frac{r_n^{(n-1)}(x)}{n!(x-x_0)} = \frac{r_n^{(n)}(x_0)}{n!} = 0,$$

т. е. действительно

$$r_n(x) = o((x-x_0)^n),\ x \to x_0.$$

Итак, доказана следующая очень важная теорема.

ТЕОРЕМА 1. *Пусть функция $f(x)$, определенная на интервале (a, b), имеет в точке $x_0 \in (a, b)$ производные до порядка n включительно. Тогда при $x \to x_0$*

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \ldots$$

$$\ldots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + o((x-x_0)^n), \tag{13.5}$$

или

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + o((x-x_0)^n).$$

Эта теорема остается справедливой вместе с ее доказательством и для функции f, определенной на отрезке $[a, b]$ при $x_0 \in [a, b]$, если для $x_0 = a$ и $x_0 = b$ под производными понимать соответствующие односторонние производные.

Формула (13.5) называется *формулой Тейлора*[1] *n-го порядка с остаточным членом в форме Пеано.*

Многочлен

$$P_n(x) = f(x_0) + \frac{f'(x_0)}{n!}(x - x_0) + \ldots + \frac{f^{(n)}(x)}{n!}(x - x_0)^n \quad (13.6)$$

называется *многочленом Тейлора степени n*, а функция

$$r_n(x) = f(x) - P_n(x) \quad (13.7)$$

— *остаточным членом n-го порядка формулы Тейлора*. Как показано, остаточный член $r_n(x)$ является бесконечно малым при $x \to x_0$ более высокого порядка, чем все члены многочлена Тейлора (13.6).

Приведем другой вид записи формулы (13.5). Положив

$$x - x_0 = \Delta x, \ \Delta y = f(x_0 + \Delta x) - f(x_0),$$

получим

$$\Delta y = \sum_{k=1}^{n} \frac{f^{(k)}(x_0)}{k!} \Delta x^k + o(\Delta x), \ \Delta x \to 0. \quad (13.5')$$

Если в формуле (13.5) $x_0 = 0$, то получается частный вид формулы Тейлора, называемый обычно *формулой Маклорена*[2]:

$$f(x) = \sum_{k=1}^{n} \frac{f^{(k)}(0)}{k!} x^k + o(x^n), \ x \to 0. \quad (13.8)$$

Доказанная теорема позволяет любую функцию, удовлетворяющую условиям этой теоремы, заменить в окрестности некоторой точки многочленом с точностью до бесконечно малых более высокого порядка, чем члены многочлена. Таким многочленом является многочлен Тейлора. Величина погрешности определяется при этом остаточным членом.

---

[1] Б. Тейлор (1685—1731) — английский математик.

[2] К. Маклорен (1698—1746) — шотландский математик.

Формула Тейлора с остаточным членом в форме Пеано дает общий метод выделения главной части функции в окрестности данной точки. На этом обстоятельстве и основаны многочисленные и разнообразные приложения формулы (13.5) в различных вопросах анализа.

Отметим полезное следствие из теоремы 1.

СЛЕДСТВИЕ. *Пусть функция $f(x)$ определена на интервале (a, b), и пусть в точке x она имеет производные до порядка $n + 1$ включительно. Тогда*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + O(x - x_0)^{n+1},\ x \to x_0. \quad (13.9)$$

Действительно, в силу теоремы 1, при $x \to x_0$

$$f(x) = \sum_{k=1}^{n+1} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + o((x - x_0)^{n+1}) \quad (13.10)$$

и, поскольку при $x \to x_0$

$$\frac{f^{(n+1)}(x_0)}{(n+1)!}(x - x_0)^{n+1} + o((x - x_0)^{n+1}) = O((x - x_0)^{n+1}),$$

из формулы (13.10) непосредственно следует формула (13.9). □

УПРАЖНЕНИЕ 1. Доказать, что если функция $f(x)$ в некоторой окрестности точки x_0 имеет производную порядка n, то, каковы бы ни были точка x этой окрестности и функция $\psi(t)$, непрерывная на отрезке с концами в точках x_0 и x, имеющая не равную нулю производную внутри этого отрезка, найдется такая точка ξ, лежащая между x_0 и x, что для остаточного члена $r_{n-1}(x)$ формулы Тейлора функции $f(x)$ имеет место формула

$$r_{n-1}(x) = \frac{\psi(x) - \psi(x_0)}{\psi'(\xi)} \frac{f^{(n)}(\xi)}{(n-1)!}(x - \xi)^{n-1},\ n = 1, 2, \ldots .$$

Получить отсюда следующие виды записи остаточного члена:

$r_{n-1}(x) = \frac{f^{(n)}(\xi)}{(n-1)!p}(x - x_0)^p(x - \xi)^{n-p}$, $p > 0$ (форма Шлёмильха—Роша[1]);

$r_{n-1}(x) = \frac{f^{(n)}(\xi)}{n!}(x - x_0)^n$ (форма Лагранжа);

$r_{n-1}(x) = \frac{f^{(n)}[x_0 + \theta(x - x_0)]}{(n-1)!}(1 - \theta)^{n-1}(x - x_0)^n$, $0 < \theta < 1$ (форма Коши).

---

[1] О. Шлёмильх (1823—1901) — немецкий математик; Э. Рош (1820—1883) — французский астроном и математик.

У к а з а н и е. Рассмотреть вспомогательную функцию

$$\varphi(t) = f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(t)}{k!}(x-t)^k$$

и применить к функциям φ и ψ теорему Коши о среднем значении. Для вывода остаточного члена в виде Шлёмильха—Роша положить $\psi(t) = (x - t)^p$.

13.2. Многочлен Тейлора как многочлен наилучшего приближения функции в окрестности данной точки

Заметим предварительно, что, очевидно, всякий многочлен

$$P_n(x) = \sum_{k=0}^{n} a_k x^k \tag{13.11}$$

может быть представлен для любого x_0 в виде

$$P_n(x) = \sum_{k=0}^{n} A_k (x - x_0)^k. \tag{13.12}$$

В самом деле, достаточно в (13.11) положить $x = x_0 + h$ и разложить правую часть по степеням h; тогда $P_n(x) = A_0 + A_1 h + \ldots + A_n h^n$, где $h = x - x_0$, т. е. получилась формула (13.12).

Докажем теперь единственность многочлена, обладающего свойством (13.2).

ТЕОРЕМА 2. *Пусть функция f дифференцируема до порядка n включительно в точке x_0, и пусть*

$$f(x) = P_n(x) + o((x - x_0)^n), \quad x \to x_0, \tag{13.13}$$

где $P_n(x) = \sum_{k=0}^{n} a_k (x - x_0)^k$ — некоторый многочлен степени, меньшей или равной n. Тогда

$$a_k = \frac{f^{(k)}(x_0)}{k!}, \quad k = 0, 1, \ldots, n, \tag{13.14}$$

т. е. $P_n(x)$ является многочленом Тейлора.

Иначе говоря, никакой многочлен степени, меньшей или равной n, отличный от многочлена Тейлора порядка n, не может приближать данную функцию с точностью до $o((x - x_0)^n)$ при $x \to x_0$ (а поэтому и с более высокой точностью $o((x - x_0)^m) m > n$, поскольку при $m > n$ имеет место со-

отношение $o((x - x_0)^m) = o((x - x_0)^n$, $x \to x_0$, — напомним, что подобные формулы читаются только слева направо). Таким образом, многочлен Тейлора является единственным многочленом, обладающим свойством (13.13), все остальные многочлены той же степени или меньшей «хуже приближают» функцию f при $x \to x_0$. Именно в этом смысле и говорят, что многочлен Тейлора является многочленом наилучшего приближения рассматриваемой функции в окрестности данной точки x_0 при $x \to x_0$.

Доказательство. Из формул (13.5) и (13.13) следует, что

$$\sum_{k=0}^{n} \frac{f^{k}(x_0)}{k!}(x - x_0)^k + o((x - x_0)^n) =$$
$$= \sum_{k=0}^{n} a_k(x - x_0)^k + o((x - x_0)^n), \quad x \to x_0,$$

откуда, перейдя к пределу при $x \to x_0$, получим $a_0 = f(x_0)$. Отбрасывая слева и справа этот член, сокращая оставшиеся в обеих частях выражения на множитель $x - x_0$ ($x \neq x_0$) и замечая, что

$$o((x - x_0)^n) = \varepsilon(x)(x - x_0)^n,$$

где $\lim\limits_{x \to x_0} \varepsilon(x) = 0$, следовательно, при $x \to x_0$ имеет место равенство

$$\frac{o((x - x_0)^n)}{x - x_0} = \varepsilon(x)(x - x_0)^{n-1} = o((x - x_0)^{n-1}),$$
$$x \neq x_0, \quad n = 1, 2, \ldots ,$$

получим

$$\sum_{k=1}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^{k-1} + o((x - x_0)^{n-1}) =$$
$$= \sum_{k=1}^{n} a_k(x - x_0)^{k-1} + o((x - x_0)^{n-1}), \ x \to x_0.$$

Переходя снова к пределу при $x \to x_0$, находим $a_1 = f'(x_0)$. Продолжая этот процесс, получаем

$$a_k = \frac{f^{(k)}(x_0)}{k!}, \ k = 0, 1, 2, \ldots , n. \ \square$$

Единственность представления функции в виде (13.13) может быть иногда использована для ее разложения по формуле Тейлора. Именно: если удается каким-либо косвенным

путем получить представление (13.13), то, в силу теоремы 2, можно утверждать, что это и есть разложение по формуле Тейлора (13.5), т. е. что коэффициенты найденного многочлена выражаются по формулам (13.14).

Так, например, соотношение (13.12) представляет собой разложение многочлена (13.11) по формуле Тейлора, причем в этом случае $r_n(x) = 0$, поэтому, в силу единственности многочлена, удовлетворяющего условию (13.13), коэффициенты многочлена (13.12) имеют вид

$$A_k = \frac{P_n^{(k)}(x_0)}{k!}.$$

Таким образом,

$$P_n(x) = \sum_{k=0}^{n} \frac{P_n^{(k)}(x_0)}{k!}(x - x_0)^k. \qquad (13.15)$$

Отметим, что из того, что многочлен (13.12) совпадает со своим многочленом Тейлора (13.15), в силу единственности представления функции в виде (13.13), следует, что если два многочлена принимают одинаковые значения на каком-нибудь интервале числовой оси, то все их коэффициенты одинаковы.

Пусть требуется разложить по формуле Тейлора функцию $f(x) = \frac{1}{1-x}$ в окрестности точки $x_0 = 0$. Замечая, что $\frac{1}{1-x}$ есть не что иное, как сумма бесконечной геометрической прогрессии $\frac{1}{1-x} = 1 + x + x^2 + \ldots + x^n + \ldots$, $|x| < 1$, и положив $r_n(x) = x^{n+1} + x^{n+2} + \ldots \frac{x^{n+1}}{1-x}$, $|x| < 1$, получим

$$\frac{1}{1-x} = 1 + x + \ldots + x^n + r_n(x),$$

где $r_n(x) = O(x^{n+1})$ и, значит, $r_n(x) = o(x^n)$ при $x \to 0$. Таким образом, представление

$$\frac{1}{1-x} = 1 + x + \ldots + x^n + o(x^n) = \sum_{k=0}^{n} x^k + o(x^n),\ x \to 0,$$

и есть разложение функции $\frac{1}{1-x}$ по формуле Тейлора в окрестности нуля.

13.3. Формулы Тейлора для основных элементарных функций

1. $f(x) = \sin x$. Функция $\sin x$ обладает производными всех порядков. Найдем для нее формулу Тейлора при $x_0 = 0$, т. е. формулу Маклорена (13.8). Было доказано (см. п. 10.1), что $(\sin x)^{(m)} = \sin\left(x + m\frac{\pi}{2}\right)$, поэтому

$$f^{(m)}(0) = \sin\frac{m\pi}{2} = \begin{cases} 0 & \text{для } m = 2k, \\ (-1)^k & \text{для } m = 2k+1, \end{cases} \quad k = 0, 1, 2, \ldots, \qquad (13.16)$$

и, согласно формуле (13.5),

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + (-1)^n\frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+2}),$$

при $x \to 0$, $n = 0, 1, 2, \ldots$, или, короче,

$$\sin x = \sum_{k=0}^{n}(-1)^k\frac{x^{2k+1}}{(2k+1)!} + o(x^{2n+2}) \text{ при } x \to 0.$$

Мы записали здесь остаточный член в виде $o(x^{2n+2})$, а не в виде $o(x^{2n+1})$, так как следующий за последним выписанным слагаемым член многочлена Тейлора, в силу (13.16), равен нулю.

2. $f(x) = \cos x$. Как известно (см. п. 10.1),

$$f^{(m)}(x) = \cos\left(x + \frac{m\pi}{2}\right),$$

поэтому

$$f^{(m)}(0) = \cos\frac{m\pi}{2} = \begin{cases} 0 & \text{для } m = 2k+1, \\ (-1)^k & \text{для } m = 2k, \end{cases} \quad k = 0, 1, 2, \ldots,$$

и

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots + (-1)^n\frac{x^{2n}}{(2n)!} + o(x^{2n+1}),$$

при $x \to 0$, или, короче,

$$\cos x = \sum_{k=0}^{n}(-1)^k\frac{x^{2k}}{(2k)!} + o(x^{2n+1})$$

при $x \to 0$, $n = 0, 1, 2, \ldots$.

3. $f(x) = e^x$. Так как $(e^x)^{(n)} = e^x$, то $f^{(n)}(0) = 1$, $n = 0, 1, \ldots$, следовательно,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!} + o(x^n), \qquad (13.17)$$

при $x \to 0$, $n = 0, 1, 2, \ldots$, или, короче,

$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + o(x^n) \text{ при } x \to 0.$$

Отсюда, заменив x через $-x$, получим

$$e^{-x} = \sum_{k=0}^{n} (-1)^k \frac{x^k}{k!} + o(x^n) \qquad (13.18)$$

при $x \to 0$, $n = 0, 1, 2, \ldots$.

4. $\operatorname{sh} x = \dfrac{e^x - e^{-x}}{2}$ и $\operatorname{ch} x = \dfrac{e^x + e^{-x}}{2}$. Сложив и вычтя (13.17) и (13.18), при $x \to 0$, $n = 0, 1, 2, \ldots$ будем иметь

$$\operatorname{sh} x = \sum_{k=0}^{n} \frac{x^{2k+1}}{(2k+1)!} + o(x^{2n+2}), \ \operatorname{ch} x = \sum_{k=0}^{n} \frac{x^{2k}}{(2k)!} + o(x^{2n+1}).$$

В силу единственности представления функции в указанном виде (см. п. 13.2), полученные соотношения являются формулами Тейлора для функций $\operatorname{sh} x$ и $\operatorname{ch} x$.

5. $f(x) = (1 + x)^{\alpha}$, где α — некоторое фиксированное число, а $x > -1$. Так как

$$f^{(n)}(x) = \alpha(\alpha - 1) \ldots (\alpha - n + 1)(1 + x)^{\alpha - n},$$

то

$$f^{(n)}(0) = \alpha(\alpha - 1) \ldots (\alpha - n + 1),$$

следовательно,

$$(1 + x)^{\alpha} = 1 + \alpha x + \frac{\alpha(\alpha - 1)}{2} x^2 +$$

$$+ \frac{\alpha(\alpha - 1)(\alpha - 2)}{3!} x^3 + \ldots + \frac{\alpha(\alpha - 1)\ldots(\alpha - n + 1)}{n!} x^n + o(x^n)$$

при $x \to 0$, или, короче,

$$(1 + x)^{\alpha} = 1 + \sum_{k=1}^{n} \frac{\alpha(\alpha - 1)\ldots(\alpha - k + 1)}{k!} x^k + o(x^n)$$

$$\text{при } x \to 0, \ n = 1, 2, \ldots . \qquad (13.19)$$

Если $\alpha = n$ — неотрицательное целое ($n = 0, 1, 2, \ldots$), то функция $f(x) = (1 + x)^n$ является многочленом степени n, и имеет место тождество

$$f(x) = Q_n(x) + o((x - x_0)^n), \ \ x \to x_0,$$

где $Q_n(x) = (1 + x)^n$ — многочлен степени n и $o((x - x_0)^n) \equiv 0$. Поэтому, в силу теоремы 2, многочлен $Q_n(x)$ является многочленом Тейлора функции $(1 + x)^n$ и, согласно (13.19),

$$(1 + x)^n = \sum_{k=0}^{n} \frac{n(n-1)\dots(n-k+1)}{k!} x^k = \sum_{k=0}^{n} C_n^k x^k.$$

Тем самым еще раз доказана формула бинома Ньютона (см. п. 2.7).

6. $f(x) = \ln(1 + x)$. Легко видеть, что

$$f'(x) = \frac{1}{1+x} = (1 + x)^{-1},\ f''(x) = -(1 + x)^{-2}$$

и, вообще, $f^{(k)}(x) = (-1)^{k-1}(k-1)!(1 + x)^{-k}$, $k = 1, 2, \dots$. Поэтому $f^{(k)}(0) = (-1)^{k-1}(k-1)!$, $k = 1, 2, \dots$, и, так как $f(0) = 0$, то

$$\ln(1 + x) = x - \frac{x^2}{2} + \dots + (-1)^{n-1}\frac{x^n}{n} + o(x^n)$$

при $x \to 0$, $n = 1, 2, \dots$, или, короче,

$$\ln(1 + x) = \sum_{k=1}^{n} (-1)^{k-1}\frac{x^k}{k} + o(x^n) \text{ при } x \to 0,\ n = 1, 2, \dots .$$

З а м е ч а н и е 1. В силу следствия теоремы 1, полученные формулы можно записать, используя символ O (O большое), следующим образом:

$$\sin x = \sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{(2k+1)!} + O(x^{2n+3}),$$

$$\cos x = \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!} + O(x^{2n+2}),$$

$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + O(x^{n+1}),$$

$$\operatorname{sh} x = \sum_{k=0}^{n} \frac{x^{2k+1}}{(2k+1)!} + O(x^{2n+3}),$$

$$\operatorname{ch} x = \sum_{k=0}^{n} \frac{x^{2k}}{(2k)!} + O(x^{2n+2}),\ \ n = 0, 1, 2, \dots ,$$

$$(1 + x)^\alpha = 1 + \sum_{k=1}^{n} \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!} x^k + O(x^{n+1}),$$

$$\ln(1 + x) = \sum_{k=1}^{n} (-1)^{k-1}\frac{x^k}{k} + O(x^{n+1}),\ \ n = 1, 2, \dots ,$$

при $x \to x_0$.

Такая запись формул Тейлора в некоторых вопросах оказывается более удобной, чем их запись с символом o (o маленькое).

З а м е ч а н и е 2. Из полученных разложений элементарных функций по формуле Тейлора в окрестности нуля можно с помощью линейной замены переменного получать и их разложения в окрестности любой точки, принадлежащей их области определения.

Например, разложим с помощью этого метода по формуле Тейлора с остаточным членом в форме Пеано функцию $f(x) = \sqrt[5]{x}$ в окрестности точки $x_0 = 1$. Положив $x = 1 + t$ и применив формулу (13.19), получим

$$\sqrt[5]{x} = (1+t)^{1/5} = 1 + \frac{1}{5}t + \frac{1}{2!}\frac{1}{5}\left(\frac{1}{5} - 1\right)t^2 + \frac{1}{3!}\frac{1}{5}\left(\frac{1}{5} - 1\right)\left(\frac{1}{5} - 2\right)t^3 +$$

$$+ \ldots + \frac{1}{n!}\frac{1}{5}\left(\frac{1}{5} - 1\right)\left(\frac{1}{5} - 2\right)\ldots\left(\frac{1}{5} - n + 1\right)t^n + o(t^n) =$$

$$= 1 + \frac{1}{5}(x-1) - \frac{1 \cdot 4}{2 \cdot 5^2}(x-1)^2 + \frac{1 \cdot 4 \cdot 9}{3!5^3}(x-1)^3 + \ldots$$

$$\ldots + \frac{(-1)^n 1 \cdot 4 \cdot 9 \ldots (5n-6)}{n!5^n}(x-1)^n + o((x-1)^n)$$

при $x \to 1$ (или, что то же самое, при $t \to 0$). Это и есть искомое разложение.

З а м е ч а н и е 3. Комбинируя указанные выше разложения функций, можно выделять главную часть (см. п. 8.4) различных элементарных функций в окрестностях тех точек, при стремлении аргумента к которым функция стремится к нулю или к бесконечности, — эти случаи встречаются наиболее часто.

В качестве примера выделим главную часть функции $\operatorname{ctg} x$ при $x \to 0$ до порядка $O(x^3)$:

$$\operatorname{ctg} x = \frac{\cos x}{\sin x} = \frac{1 - \frac{x^2}{2} + O(x^4)}{x - \frac{x^3}{6} + O(x^5)} =$$

$$= \frac{1}{x}\left[1 - \frac{x^2}{2} + O(x^4)\right]\left[1 - \frac{x^2}{6} + O(x^4)\right]^{-1} =$$

$$= \frac{1}{x}\left[1 - \frac{x^2}{2} + O(x^4)\right]\left[1 + \frac{x^2}{6} - O(x^4) + O\left(\left(-\frac{x^2}{6} + O(x^4)\right)^2\right)\right] =$$

$$= \frac{1}{x}\left[1 + \frac{x^2}{6} - \frac{x^2}{2} + O(x^4)\right] = \frac{1}{x} - \frac{1}{3}x + O(x^3), \quad x \to 0.$$

Здесь были использованы разложения по формуле Тейлора:

косинуса

$$\cos x = 1 - \frac{x^2}{2} + O(x^4),\ \ x \to 0,$$

синуса

$$\sin x = x - \frac{x^3}{6} + O(x^5),\ \ x \to 0,$$

бинома

$$(1 + u)^{\alpha} = 1 + \alpha u + O(u^2)$$

при $\alpha = -1$ и $u = -\frac{x^2}{6} + O(x^4),\ \ x \to 0.$

13.4. Вычисление пределов с помощью формулы Тейлора (метод выделения главной части)

Формула Тейлора дает простое и весьма общее правило для выделения главной части функции. В результате этого метод вычисления пределов функций с помощью выделения главной части приобретает законченный алгоритмический характер.

Рассмотрим сначала случай неопределенности вида $\frac{0}{0}$. Пусть требуется найти предел $\lim\limits_{x \to x_0} \frac{f(x)}{g(x)}$, где $\lim\limits_{x \to x_0} f(x) = \lim\limits_{x \to x_0} g(x) = 0$. В этом случае рекомендуется разложить по формуле Тейлора функции f и g в окрестности точки x_0 (если, конечно, это возможно), ограничившись в этом разложении лишь первыми не равными нулю членами, т. е. взять разложения в виде

$$f(x) = a(x - x_0)^n + o((x - x_0)^n),\ a \neq 0,$$

$$g(x) = b(x - x_0)^m + o((x - x_0)^m),\ b \neq 0,$$

тогда

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{a(x - x_0)^n + o((x - x_0)^n)}{b(x - x_0)^m + o((x - x_0)^m)} =$$

$$= \frac{a}{b} \lim_{x \to x_0} (x - x_0)^{n - m} = \begin{cases} 0, & \text{если } n > m, \\ \frac{a}{b}, & \text{если } n = m, \\ \infty, & \text{если } n < m. \end{cases}$$

Часто бывает удобно для разложения функций f и g по формуле Тейлора использовать готовый набор разложений элементарных функций, полученный в п. 13.3. Для этого следует в случае $x_0 \neq 0$ предварительно выполнить замену переменного $t = x - x_0$; тогда $x \to x_0$ будет соответствовать $t \to 0$. Случай $x \to \infty$ заменой переменного $x = 1/t$ сводится к случаю $t \to 0$.

Если имеется неопределенность вида $\frac{\infty}{\infty}$, т. е. требуется найти $\lim\limits_{x \to x_0} \frac{f(x)}{g(x)}$, где $\lim\limits_{x \to x_0} f(x) = \lim\limits_{x \to x_0} g(x) = \infty$, то ее легко привести к рассмотренному случаю $\frac{0}{0}$ преобразованием $\frac{f(x)}{g(x)} = \frac{1/g(x)}{1/f(x)}$.

Подобно вычислению пределов с помощью правила Лопиталя, при применении метода выделения главной части к раскрытию неопределенностей вида $0 \cdot \infty$ и $\infty - \infty$ их следует преобразовать к неопределенности вида $\frac{0}{0}$. Наконец, для раскрытия неопределенностей вида 0^0, ∞^0 и 1^∞ указанным методом необходимо предварительно прологарифмировать рассматриваемые функции. Посмотрим на примерах, как применяется формула Тейлора к вычислению пределов функций. Пусть требуется найти

$$\lim_{x \to x_0} \frac{e^x - e^{-x} - 2x}{x - \sin x}.$$

Заметив, что (см. п. 13.3)

$$e^x = 1 + x + x^2/2 + x^3/3 + o(x^3),$$
$$e^{-x} = 1 - x + x^2/2 - x^3/3 + o(x^3),$$
$$\sin x = x - x^3/3! + o(x^3),$$

получим

$$\lim_{x \to 0} \frac{e^x - e^{-x} - 2x}{x - \sin x} =$$
$$= \lim_{x \to 0} \frac{x^3/3 + o(x^3)}{x^3/6 + o(x^3)} = \lim_{x \to 0} \frac{x^3/3}{x^3/6} = 2.$$

Рассмотрим неопределенность вида $\infty - \infty$:

$$\lim_{x \to 0}\left(\frac{1}{x^2} - \frac{1}{\sin^2 x}\right) = \lim_{x \to 0}\frac{\sin^2 x - x^2}{x^2 \sin^2 x} =$$

$$= \lim_{x \to 0}\frac{[x - x^3/6 + o(x^3)]^2 - x^2}{x^2[x + o(x)]^2} = \lim_{x \to 0}\frac{-x^4/3 + o(x^4)}{x^2[x^2 + o(x^2)]} =$$

$$= \lim_{x \to 0}\frac{-x^4/3 + o(x^4)}{x^4 + o(x^4)} = \lim_{x \to 0}\frac{-x^4/3}{x^4} = -\frac{1}{3}.$$

В качестве последнего примера вычислим предел $\lim\limits_{x \to 0}\left(\frac{\sin x}{x}\right)^{1/x}$, т. е. раскроем неопределенность вида 1^{∞}. Согласно общему правилу, найдем предел логарифма выражения, стоящего под знаком предела:

$$\lim_{x \to 0}\frac{1}{x}\ln\frac{\sin x}{x} = \lim_{x \to 0}\frac{\ln\frac{x + o(x^2)}{x}}{x} =$$

$$= \lim_{x \to 0}\frac{\ln(1 + o(x))}{x} = \lim_{x \to 0}\frac{o(x)}{x} = 0.$$

Следовательно,

$$\lim_{x \to 0}\left(\frac{\sin x}{x}\right)^{\frac{1}{x}} = e^{\lim\limits_{x \to 0}\frac{1}{x}\ln\frac{\sin x}{x}} = 1.$$

§ 14.

Исследование поведения функций

14.1. Признак монотонности функции

Опишем свойство монотонности функции с помощью ее производной.

ТЕОРЕМА 1. *Для того чтобы непрерывная на некотором промежутке функция, дифференцируемая во всех его внутренних точках, возрастала (убывала) на этом промежутке, необходимо и достаточно, чтобы производная функции была во всех внутренних точках промежутка неотрицательна (неположительна).*

Если во всех внутренних точках промежутка производная функции положительна (отрицательна), то функция строго возрастает (строго убывает) на этом промежутке.

Доказательство необходимости. Если функция f возрастает (убывает) на промежутке Δ (отрезке, интервале или полуинтервале) с концами в точках a и b, если $x_0 \in \Delta$, $\Delta x > 0$, $x_0 + \Delta x \in \Delta$, то $f(x_0 + \Delta x) \geqslant f(x_0)$ (соответственно $f(x_0 + \Delta x) \leqslant f(x_0)$), поэтому $\Delta y = f(x_0 + \Delta x) - f(x_0) \geqslant 0$ (соответственно $\Delta y \leqslant 0$).

Следовательно, $\frac{\Delta y}{\Delta x} \geqslant 0$ $\left(\text{соответственно } \frac{\Delta y}{\Delta x} \leqslant 0\right)$. Перейдя к пределу при $\Delta x \to 0$, получим $f'(x_0) \geqslant 0$ (соответственно $f'(x_0) \leqslant 0$).

Доказательство достаточности. Пусть $x_1 < x_2$, $x_1 \in \Delta$, $x_2 \in \Delta$, тогда по формуле Лагранжа (см. п. 11.2), $f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1)$, где $x_1 < x_2$. Так как $x_2 - x_1 > 0$, то при $f'(x) \geqslant 0$ на (a, b) (откуда следует, что, в частности, $f'(\xi) \geqslant 0$) будем иметь $f(x_2) \geqslant f(x_1)$, т. е. функция f возрастает. Аналогично, при $f'(x) \leqslant 0$ на (a, b) имеем $f'(\xi) \leqslant 0$ и, следовательно, $f(x_2) \leqslant f(x_1)$, т. е. функция f убывает.

Если $f'(x) > 0$ на (a, b), то $f'(\xi) > 0$ и поэтому $f(x_2) > f(x_1)$, т. е. функция f строго возрастает. Если же $f'(x) < 0$ на (a, b), то $f'(\xi) < 0$, следовательно, $f(x_2) < f(x_1)$, т. е. функция f строго убывает. □

Отметим, что условия $f'(x) > 0$ и $f'(x) < 0$ не являются необходимыми для строгого возрастания (строгого убывания) дифференцируемой на интервале функции, что показывают примеры функций $f_1(x) = x^3$ и $f_2(x) = -x^3$. Первая из них строго возрастает, а вторая строго убывает на всей числовой оси, но при $x = 0$ их производные обращаются в нуль.

Аналогичная теорема верна для непрерывных функций, не имеющих в конечном числе точек производной. Утверждение второй части теоремы остается в силе, если, кроме того, в конечном числе точек производная обращается в нуль. Например,

если функция непрерывна на некотором интервале и имеет всюду в нем положительную (отрицательную) про-

изводную, кроме, быть может, конечного числа точек, в которых производная обращается в нуль или не существует, то функция строго возрастает (строго убывает) на рассматриваемом интервале.

Это непосредственно следует из теоремы 1: достаточно ее последовательно применить ко всем промежуткам, на которые разбивается заданный интервал указанным конечным множеством точек.

Пример. Исследуем функцию

$$f(x) \stackrel{\text{def}}{=} \begin{cases} \dfrac{\sin x}{x} & \text{при } 0 < x \leqslant \dfrac{\pi}{2}, \\ 1 & \text{при } x = 0. \end{cases}$$

Функция f дифференцируема (следовательно, и непрерывна) на отрезке $[0, \pi/2]$. В специальной проверке нуждается лишь существование производной в точке $x = 0$. Применяя, например, дважды правило Лопиталя, получаем

$$\lim_{\Delta x \to 0} \frac{f(\Delta x) - f(0)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\dfrac{\sin \Delta x}{\Delta x} - 1}{\Delta x} =$$

$$= \lim_{t \to 0} \frac{\sin t - t}{t^2} = \lim_{t \to 0} \frac{\cos t - 1}{2t} = -\frac{1}{2} \lim_{t \to 0} \sin t = 0.$$

Это и означает, что существует $f'(0) = 0$.

Для всех $x \neq 0$ имеем

$$f'(x) = \left(\frac{\sin x}{x}\right)' = \frac{x \cos x - \sin x}{x^2} = \frac{\cos x}{x^2}(x - \operatorname{tg} x) < 0,$$

так как $x < \operatorname{tg} x$, если $0 < x < \pi/2$ (см. доказательство леммы 1 в п. 8.1). Следовательно, функция f строго убывает на отрезке $[0, \pi/2]$, поэтому $f(0) > f(x) > f(1)$, т. е.

$$\frac{2}{\pi} < \frac{\sin x}{x} < 1 \text{ при } 0 < x < \frac{\pi}{2}. \qquad (14.1)$$

УПРАЖНЕНИЕ 1. Доказать, что, для того чтобы дифференцируемая на интервале (a, b) функция строго возрастала на нем, необходимо и достаточно, чтобы во всех точках $x \in (a, b)$ выполнялось неравенство $f'(x) \geqslant 0$ и чтобы для любых точек $x' \in (a, b)$ и $x'' \in (a, b)$, $x' < x''$, существовала бы точка $\xi \in (x', x'')$, для которой $f'(\xi) > 0$.

14.2. Отыскание наибольших и наименьших значений функции

Определение 1. *Пусть функция f определена в некоторой окрестности точки x_0. Тогда x_0 называется точкой максимума (точкой минимума) функции f, если существует такое $\delta > 0$, что для всех Δx, удовлетворяющих условию $|\Delta x| < \delta$, выполняется неравенство $f(x_0 + \Delta x) \leqslant f(x_0)$ (неравенство $f(x_0 + \Delta x) \geqslant f(x_0)$).*

Если существует такое $\delta > 0$, что для всех $\Delta x \neq 0$ таких, что $|\Delta x| < \delta$, выполняется неравенство $f(x_0 + \Delta x) < f(x_0)$ (неравенство $f(x_0 + \Delta x) > f(x_0)$), то x_0 называется точкой строгого максимума (строгого минимума).

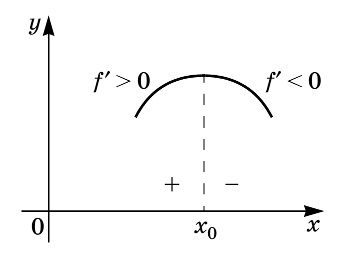
Точки максимума (строгого максимума) и минимума (строгого минимума) называются *точками экстремума* (*строгого экстремума*).

Для точек x_0 строгого экстремума функции f, и только для них, приращение $\Delta f = f(x_0 + \Delta x) - f(x_0)$ не меняет знака при переходе аргумента через x_0, т. е. при изменении знака Δx. Именно: $\Delta f < 0$ для точек строгого максимума и $\Delta f > 0$ в случае строгого минимума независимо от знака достаточно малого $\Delta x \neq 0$.

ТЕОРЕМА 2 (необходимые условия экстремума). *Пусть x_0 является точкой экстремума функции f, определенной в некоторой окрестности точки x_0. Тогда либо производная $f'(x)$ не существует, либо $f'(x_0) = 0$.*

Доказательство. Действительно, если x_0 является точкой экстремума для функции f, то найдется такая окрестность $U(x_0, \delta)$, что значение функции f в точке x_0 будет наибольшим или наименьшим на этой окрестности. Поэтому если в точке x_0 существует производная, то она, согласно теореме Ферма (см. п. 11.1), равна нулю. □

Отметим, что условие $f'(x_0) = 0$ не является (для дифференцируемой при $x = x_0$ функции) достаточным условием наличия экстремума, как это показывает пример функции $f(x) = x^3$, которая при $x = 0$ имеет производную, равную нулю, но для которой $x = 0$ не является точкой экстремума.



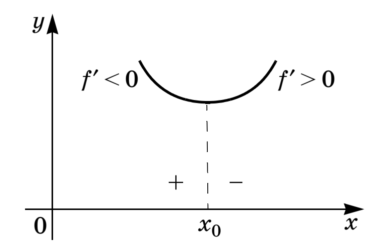


Рис. 57

УПРАЖНЕНИЕ 2 (достаточные условия экстремума). Пусть функция f определена на интервале (a, b) и непрерывна в точке $x_0 \in (a, b)$. Доказать, если f (строго) возрастает на интервале (a, x_0) и (строго) убывает на (x_0, b), то x_0 является точкой (строгого) максимума; если же функция f (строго) убывает на (a, x_0) и (строго) возрастает на (x_0, b), то x_0 является точкой (строгого) минимума.

ТЕОРЕМА 3 (достаточные условия строгого экстремума). *Пусть функция f дифференцируема в некоторой окрестности точки x_0, кроме, быть может, самой точки $x_0 \in (a, b)$, в которой она является, однако, непрерывной. Если производная $f'(x)$ меняет знак при переходе через x_0 (это означает, что существует такое число $\delta > 0$, что значения производной f' имеют один и тот же знак всюду в $(x_0 - \delta, x_0)$ и противоположный знак для всех $x \in (x_0, x_0 + \delta)$), то x_0 является точкой строгого экстремума.*

При этом если при $x_0 - \delta < x < x_0$ выполняется неравенство $f'(x) > 0$, а при $x_0 < x < x_0 + \delta$ — неравенство $f'(x) < 0$, то x_0 является точкой строгого максимума, а если при $x_0 - \delta < x < x_0$ выполняется неравенство $f'(x) < 0$, а при $x_0 < x < x_0 + \delta$ — неравенство $f'(x) > 0$, то x_0 является точкой строгого минимума (рис. 57).

Доказательство. Рассмотрим случай $f'(x) > 0$ для $x < x_0$ и $f'(x) < 0$ для $x > x_0$, где x принадлежит окрестности точки x_0, указанной в условиях теоремы. По теореме Лагранжа (см. п. 11.2),

$$\Delta f = f(x) - f(x_0) = f'(\xi)(x - x_0),$$

где ξ лежит на интервале с концами x_0 и x.

Если $x < x_0$, то $x - x_0 < 0$ и $f'(\xi) > 0$, так как $x < \xi < x_0$. Если $x > x_0$, то $x - x_0 > 0$ и $f'(\xi) < 0$, так как в этом случае $x_0 < \xi < x$. Таким образом, всегда $\Delta f < 0$, т. е. точка x_0 является точкой строгого максимума. Аналогично рассматривается второй случай. □

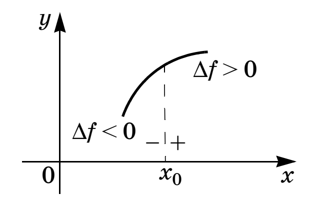
Из п. 14.1 следует, что если функция имеет всюду в некоторой проколотой окрестности данной точки x_0 производную одного и того же знака, а в самой точке x_0 производная либо равна нулю, либо не существует, однако сама функция непрерывна, т. е. если производная непрерывной функции «не меняет знака» при переходе через точку x_0, то эта точка заведомо *не является точкой экстремума* рассматриваемой функции (более того, функция в указанной окрестности строго возрастает или строго убывает в зависимости от того, положительна или отрицательна производная в точках $x \neq x_0$).

Объединяя это утверждение с доказанной выше теоремой 3, получим следующий результат.

Если функция $f(x)$ определена в некоторой окрестности точки x_0, непрерывна при $x = x_0$, имеет всюду в рассматриваемой окрестности, кроме, может быть, точки x_0, производную и эта производная с каждой стороны от x_0 сохраняет постоянный знак (следовательно, можно говорить о сохранении или перемене знака у производной при переходе через x_0), то, для того чтобы при $x = x_0$ функция достигла экстремума, необходимо и достаточно, чтобы производная меняла знак при переходе через точку x_0.

Следует, однако, обратить внимание на то, что рассмотренный здесь случай, т. е. случай, когда можно в указанном смысле говорить о перемене знака производной при переходе через точку x_0, не исчерпывает возможные ситуации (даже для всюду дифференцируемых функций): может случиться, что в сколь угодно малых односторонних окрестностях точки x_0 производная функции меняет знак. В этом случае приходится применять другие методы для исследования функции на экстремум при $x = x_0$.

Таким образом, в более широком классе функций, дифференцируемых в окрестности рассматриваемой точки, кроме, быть может, самой этой точки, условие изменения знака



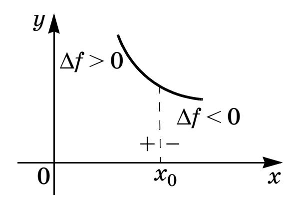


Рис. 58

производной в точке является лишь достаточным условием экстремума.

Задача 10. Построить пример функции, которая дифференцируема на интервале, достигает в некоторой его точке x_0 строгого экстремума, а ее производная в любой окрестности точки (как слева, так и справа от нее) принимает и положительные, и отрицательные значения (таким образом, показать, что условие изменения знака производной в данной точке, являясь достаточным для наличия строгого экстремума, не является вместе с тем необходимым).

Введем еще одно понятие, которым будем пользоваться в дальнейшем.

Определение 2. *Пусть функция f определена в некоторой окрестности точки x_0. Будем называть x_0 точкой возрастания (убывания) функции f, если существует такое $\delta > 0$, что при $x_0 - \delta < x < x_0$ выполняется неравенство $f(x) < f(x_0)$ (неравенство $f(x) > f(x_0)$), а при $x_0 < x < x_0 + \delta$ — неравенство $f(x) > f(x_0)$ (неравенство $f(x) < f(x_0)$).*

Таким образом, точки возрастания и убывания функции f характеризуются тем, что при переходе через них приращение функции меняет знак с «−» на «+» в точке возрастания и с «+» на «−» в точке убывания (рис. 58).

Не следует думать, что если функция определена на интервале, то всякая точка этого интервала является либо точкой экстремума функции, либо точкой возрастания, либо точкой убывания: могут существовать точки, не принадлежащие ни к одному из указанных типов. Например, точка $x = 0$ для функции

$$y = \begin{cases} x^2 \sin \frac{1}{x} & \text{при } x \neq 0, \\ 0 & \text{при } x = 0 \end{cases} \tag{14.2}$$

не является ни точкой экстремума, ни точкой возрастания, ни точкой убывания.

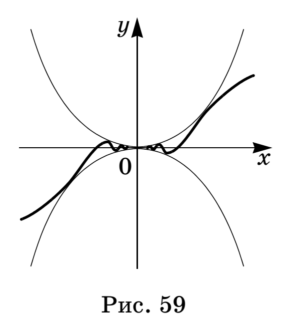


Рис. 59

Производная функции (14.2) равна (см. пример 8 в п. 9.7)

$$y' = \begin{cases} 2x \sin \frac{1}{x} - \cos \frac{1}{x} & \text{при } x \neq 0, \\ 0 & \text{при } x = 0. \end{cases} \quad (14.3)$$

Таким образом, функция (14.2) дифференцируема на всей числовой оси. При $x = 0$ ее производная имеет разрыв второго рода, поскольку

$$\lim_{x \to 0} 2x \sin \frac{1}{x} = 0, \quad (14.4)$$

а второе слагаемое в выражении (14.3) для производной y' при $x \neq 0$, т. е. $-\cos \frac{1}{x}$ не имеет предела при $x \to 0$. Кроме того, это слагаемое, изменяясь в любой односторонней окрестности точки $x = 0$ от -1 до $+1$, бесконечно много раз меняет в ней знак. Отсюда, в силу формул (14.3) и (14.4), следует, что и производная функции (14.2) в любой сколько угодно малой односторонней окрестности нуля также меняет знак. Общий характер поведения функции (14.2) изображен на рисунке 59.

Сформулируем теперь основанные на использовании производных высших порядков достаточные условия наличия строгого экстремума, а также точек возрастания и убывания.

ТЕОРЕМА 4. *Пусть в точке x_0 у функции f существуют производные до порядка $n \geqslant 1$ включительно, причем*

$$f^{(i)}(x_0) = 0 \ \ \text{для } i = 1, 2, \ldots, n - 1, \ \ f^{(n)}(x_0) \neq 0. \quad (14.5)$$

Тогда если $n = 2k$, $k = 1, 2, \ldots$, т. е. n — четное число, то функция f имеет в точке x_0 строгий экстремум, а именно максимум при $f^{(2k)}(x_0) < 0$ и минимум при $f^{(2k)}(x_0) > 0$. Если же $n = 2k + 1$, $k = 0, 1, \ldots$, т. е. n — нечетное число, то функция f не имеет в точке x_0 экстремума; в этом случае x_0 является точкой возрастания при $f^{(2k+1)}(x_0) > 0$ и убывания при $f^{(2k+1)}(x_0) < 0$.

Предпошлем доказательству теоремы одно простое замечание: если в некоторой окрестности точки x_0 имеет место соотношение $\beta(x) = o\,(\alpha(x))$, $x \to x_0$, то существует такое $\delta > 0$, что при $|x - x_0| < \delta$ справедливо неравенство

$$|\beta(x)| \leqslant \frac{1}{2}|\alpha(x)|.$$

В самом деле,

$$\beta(x) = \varepsilon(x)\alpha(x),$$

где $\lim\limits_{x \to x_0} \varepsilon(x) = 0$, поэтому существует такое $\delta > 0$, что при $|x - x_0| < \delta$ выполняется неравенство $|\varepsilon(x)| < \frac{1}{2}$. Отсюда и следует указанное выше неравенство.

Доказательство. Прежде всего заметим, что функция f имеет в точке x_0 производную порядка $n \geqslant 1$, поэтому (согласно определению производной) производная порядка $n - 1$ рассматриваемой функции определена в некоторой окрестности точки x_0. Следовательно, и сама функция f также определена, во всяком случае в той же окрестности точки x_0.

Запишем формулу Тейлора n-го порядка для функции f в окрестности точки x_0. В силу формулы (13.5) и условий (14.5), имеем

$$\Delta f = f(x_0 + \Delta x) - f(x_0) = \frac{f^{(n)}(x_0)}{n!}\Delta x^n + \alpha(x), \qquad (14.6)$$

где $\alpha(x) = o(\Delta x^n)$[1], $\Delta x \to 0$, откуда, в силу условия $f^{(n)}(x_0) \neq 0$, вытекает (см. п. 8.2), что

$$\alpha(x) = o\left(\frac{f^{(n)}(x_0)}{n!}\Delta x^n\right), \quad \Delta x \to 0.$$

Поэтому, согласно сделанному замечанию, существует такое $\delta > 0$, что при $|\Delta x| < \delta$

$$|\alpha(x)| < \frac{1}{2}\left|\frac{f^{(n)}(x_0)}{n!}\Delta x^n\right|.$$

Из этого неравенства следует, что при $|\Delta x| < \delta$, $\Delta x \neq 0$, знак правой части равенства (14.6), следовательно, и знак Δf совпадают со знаком первого слагаемого правой части.

Если $n = 2k$, $k = 1, 2, \dots$, то в (14.6) приращение аргумента Δx возводится в четную степень, поэтому знак приращения функции Δf не зависит от знака $\Delta x \neq 0$ и, значит, x_0 является точкой строгого экстремума, причем точкой строгого макси-

---

[1] Здесь $\Delta x^n \stackrel{\text{def}}{=} (\Delta x)^n$ (а не $\Delta(x^n)$).

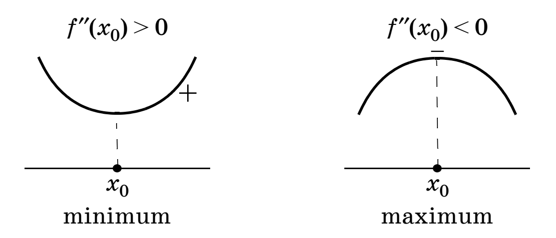


Рис. 60

мума при $f^{(2k)}(x_0) < 0$ (в этом случае $\Delta f < 0$) и строгого минимума при $f^{(2k)}(x_0) > 0$ (в этом случае $\Delta f > 0$).

Если же $n = 2k + 1$, $k = 0, 1, 2, \ldots$, то Δx возводится в нечетную степень, поэтому знак Δf меняется вместе с изменением знака Δx и, значит, x_0 не является точкой экстремума. Если Δx меняет знак с «−» на «+», то при $f^{(2k+1)}(x_0) > 0$ приращение Δf также меняет знак с «−» на «+» и, значит, x_0 является точкой возрастания функции f, а при $f^{(2k+1)}(x_0) < 0$ приращение Δf меняет знак с «+» на «−» и, значит, x_0 является точкой убывания функции f. □

Из доказанной теоремы вытекают, в частности, при $n = 1$ и $n = 2$ два следствия.

1. *Если* $f'(x_0) > 0$, *то* x_0 *является точкой возрастания функции*; *если* $f'(x_0) < 0$, то x_0 — *точка убывания функции.*

2. *Если* $f'(x_0) = 0$, *а* $f''(x_0) \neq 0$, *то при* $f''(x_0) > 0$ x_0 *является точкой строгого минимума, а при* $f''(x_0) < 0$ — *точкой строгого максимума функции* f (рис. 60).

Следствие 1 остается в силе и для бесконечных производных со знаком: если $f'(x_0) = +\infty$ (соответственно $f'(x_0) = -\infty$), то x_0 является точкой возрастания (убывания) функции. В самом деле, если, например, $f'(x_0) = +\infty$, то для любого $\varepsilon > 0$ и, в частности, для $\varepsilon = 1$ существует такое $\delta > 0$, что при всех Δx, удовлетворяющих условию $|\Delta x| < \delta$, имеет место неравенство $\frac{\Delta y}{\Delta x} > 1$. Поэтому при $0 < \Delta x < \delta$ имеем $\Delta y > \Delta x > 0$, а при $-\delta < \Delta x < 0$ аналогично имеем $\Delta y < \Delta x < 0$, т. е. x_0 — точка возрастания. Аналогично рассматривается случай, когда $f'(x_0) = -\infty$.

Заметим, что из первого следствия еще раз вытекает теорема Ферма (см. теорему 1 п. 11.1). Действительно, если функция $f(x)$ определена в некоторой окрестности точки x_0 и имеет в этой точке экстремум, то производная в x_0 не может быть ни положительной, ни отрицательной, так как в противном случае функция либо возрастала, либо убывала бы в этой точке. Следовательно, производная в x_0 или не существует, или, если существует, необходимо равна нулю.

Отметим еще, что из теоремы 4 непосредственно вытекает следующий критерий наличия точек экстремума.

Пусть у функции f в точке x_0 существуют производные до порядка $n \geqslant 1$ включительно, причем

$$f^{(k)}(x_0) = 0,\ \ k = 1, 2, \ldots, n - 1,\ \ f^{(n)}(x_0) \neq 0.$$

Тогда, для того чтобы при $x = x_0$ функция достигала экстремума, необходимо и достаточно, чтобы n было четным числом.

Все полученные правила справедливы лишь в том случае, когда функция f определена в некоторой окрестности точки x_0. Однако об экстремуме функции можно говорить не только в этом случае: пусть функция f определена на некотором числовом множестве X; будем называть $x_0 \in X$ *точкой максимума* (*минимума*)[1], *если существует такое* $\delta > 0$, что если $x \in X$ и $|x - x_0| < \delta$, то $f(x) \leqslant f(x_0)$ (соответственно $f(x) \geqslant f(x_0)$). Подобным же образом определяются в этом случае и понятия строгого максимума и строгого минимума, следует лишь при $x \neq x_0$ знаки нестрогих неравенств заменить знаками строгих неравенств.

Например, если функция f определена на полуинтервале $[a, b)$, то точка a в указанном смысле может являться экстремальной. Заметим, однако, что производная (правосторонняя) в этой точке, вообще говоря, не обязана обращаться в нуль (см. п. 11.1). Так, функция $y = x$ на отрезке $[0, 1]$ имеет строгий минимум при $x = 0$ и строгий максимум при $x = 1$, однако в этих точках, как и всюду на отрезке $[0, 1]$, $y' = 1$.

---

[1] Правильнее было бы добавить — локального, но не будем усложнять терминологию.

Выяснение обстоятельства, имеет или нет функция экстремум на концах промежутка, принадлежащего ее области определения (такой экстремум будем называть концевым), требует специального исследования.

УПРАЖНЕНИЕ 3. Пусть функция f определена на отрезке $[a, b]$ и имеет производные при $x = a$ и $x = b$. Доказать, что если $f'_+(a) > 0$ (соответственно $f'_-(b) < 0$), то точка $x = a$ (соответственно $x = b$) является точкой строгого минимума, а если $f'_+(a) < 0$ (соответственно $f'_-(b) > 0$), то $x = a$ (соответственно $x = b$) является точкой строгого максимума.

Установленные теоремы лежат в основе метода, позволяющего единообразно решать многочисленные математические, физические, экономические и технические задачи, в которых изучаются экстремальные значения какой-либо величины.

Пусть, например, требуется определить наибольшее значение функции f на отрезке $[a, b]$. Может случиться, что это возможно сделать достаточно просто каким-либо способом исходя из конкретного вида функции. Если же не ясно, как это сделать, то следует найти все ее критические точки, лежащие на $[a, b]$ (точка, в которой функция определена, а ее производная либо равна нулю, либо не существует, обычно называется *критической точкой* этой функции). Затем из этих значений x необходимо исходя из сказанного отобрать те, в которых возможен максимум (можно заведомо отбросить точки, удовлетворяющие достаточным условиям наличия минимума). После этого достаточно сравнить между собой по величине значения функции в полученных точках и числа $f(a)$ и $f(b)$; наибольшее из этих чисел и является наибольшим значением функции на отрезке $[a, b]$. Эта задача принципиально заведомо может быть решена, если множество критических точек конечно.

Если функция определена на полуинтервале (конечном или бесконечном), например на полуинтервале вида $[a, b)$, задача об определении ее наибольшего значения на этом полуинтервале требует дополнительных исследований; найдя множество указанных выше точек, надо изучить еще поведе-

ние функции при $x \to b - 0$. Аналогично решаются и задачи на определение наименьших значений функций.

Не следует, однако, думать, что изложенный метод позволяет находить точки экстремума данной функции с любой нужной степенью точности. Это не так, поскольку если его использовать, надо прежде всего уметь решать уравнение $f'(x) = 0$ с заданной степенью точности, что является другой математической задачей. Как она решается с помощью дифференциального исчисления в тех случаях, когда точное решение уравнения не записывается в явном виде, будет показано в дальнейшем (см. § 62).

Пример. Две точки движутся с постоянными скоростями v_1 и v_2 по двум прямым, образующим прямой угол, в направлении вершины этого угла, от которой в начале движения первая точка находилась на расстоянии a, а вторая — на расстоянии b. Через какое время после начала движения расстояние между точками будет наименьшим?

Пусть $\rho = \rho(t)$ — расстояние между точками через время t после начала движения, которое будем считать начавшимся при $t = 0$. Тогда

$$\rho^2(t) = (a - v_1 t)^2 + (b - v_2 t)^2.$$

Функция $\rho(t)$, очевидно, достигает минимума при том же значении t, при котором достигает минимума функция $y = \rho^2(t)$.

Из физических соображений ясно, что расстояние $\rho(t)$ должно достигать минимума (тела начинают сближаться), а максимума заведомо нет, ибо $\rho(t) \to +\infty$ при $t \to +\infty$. В силу необходимого условия экстремума, это может быть только в точке, в которой $y' = 0$, и так как $y' = -2v_1(a - v_1 t) - 2v_2(b - v_2 t)$, то из условия $y' = 0$ получаем единственное значение $t_0 = \frac{av_1 + bv_2}{v_1^2 + v_2^2}$, которое и дает ответ на поставленный вопрос.

14.3. Выпуклость и точки перегиба

Пусть функция f определена на интервале (a, b) и пусть $a < x_1 < x_2 < b$. Проведем прямую через точки $A = (x_1, f(x_1))$

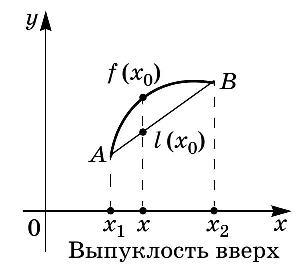




Рис. 61

и $B = (x_2, f(x_2))$, лежащие на графике функции f. Ее уравнением является

$$y = \frac{f(x_2)(x - x_1) + f(x_1)(x_2 - x)}{x_2 - x_1}.$$

Обозначим правую часть этого уравнения через $l(x)$; тогда оно кратко запишется в виде

$$y = l(x).$$

Очевидно, $l(x_1) = f(x_1)$, $l(x_2) = f(x_2)$.

Определение 3. *Функция f называется выпуклой вверх (выпуклой вниз) на интервале (a, b), если, каковы бы ни были точки x_1 и x_2, $a < x_1 < x_2 < b$, для любой точки x_0 интервала (x_1, x_2) выполняется неравенство*

$$l(x_0) \leqslant f(x_0) \tag{14.7}$$

(*соответственно неравенство*

$$l(x_0) \geqslant f(x_0)). \tag{14.8}$$

Геометрически это означает, что любая точка хорды AB (т. е. отрезка прямой $y = l(x)$ с концами в точках A и B) лежит не выше (не ниже) точки графика функции f, соответствующей тому же значению аргумента (рис. 61).

Определение 4. *Если вместо* (14.7) *и* (14.8) *выполняются строгие неравенства $l(x_0) < f(x_0)$ и соответственно $l(x_0) > f(x_0)$ при любых x_0, x_1 и x_2 таких, что $a < x_1 < x_0 < x_2 < b$, то функция f называется строго выпуклой вверх (строго выпуклой вниз) на интервале (a, b).*

В этом случае любая точка хорды AB, исключая ее концы, лежит ниже (выше) соответствующей точки графика функции.

Определение 5. *Всякий интервал, на котором функция (строго) выпукла вверх, соответственно вниз, называется интервалом (строгой) выпуклости вверх, соответственно вниз, этой функции.*

ТЕОРЕМА 5 (достаточное условие строгой выпуклости). *Пусть функция f дважды дифференцируема на интервале (a, b). Тогда, если $f'' < 0$ на (a, b), то функция f строго выпукла вверх, а если $f'' > 0$ на (a, b), то функция f строго выпукла вниз на этом интервале.*

Доказательство. Пусть $a < x_1 < x_2 < b$. Тогда

$$l(x) - f(x) = \frac{f(x_2)(x - x_1) + f(x_1)(x_2 - x)}{x_2 - x_1} - f(x)\frac{(x - x_1) + (x_2 - x)}{x_2 - x_1} =$$

$$= \frac{[f(x_2) - f(x)](x - x_1) - [f(x) - f(x_1)](x_2 - x)}{x_2 - x_1}.$$

Применяя теорему Лагранжа (см. п. 11.2), получаем

$$l(x) - f(x) = \frac{f'(\eta)(x_2 - x)(x - x_1) - f'(\xi)(x - x_1)(x_2 - x)}{x_2 - x_1} =$$

$$= \frac{[f'(\eta) - f'(\xi)](x_2 - x)(x - x_1)}{x_2 - x_1},$$

где $x_1 < \xi < x < \eta < x_2$.

Воспользуемся снова теоремой Лагранжа:

$$l(x) - f(x) = \frac{f''(\zeta)(x_2 - x)(x - x_1)(\eta - \xi)}{x_2 - x_1}, \ \xi < \zeta < \eta.$$

Отсюда видно, что если $f'' < 0$ на (a, b), следовательно, в частности, $f''(\zeta) < 0$, то $l(x) < f(x)$, т. е. функция f строго выпукла вверх; если же $f'' > 0$ на (a, b), то $l(x) > f(x)$, т. е. функция f строго выпукла вниз. □

Замечание 1. Из доказательства теоремы 5 видно, что если условие положительности второй производной на интервале заменить условием ее неотрицательности, то функция будет выпукла вниз (вообще говоря, нестрого) на этом интервале. Соответственно, если вторая производная неположительна на интервале, то функция выпукла вверх на этом интервале.

Условие знакопостоянства второй производной, являясь достаточным для строгой выпуклости (вверх или вниз), не

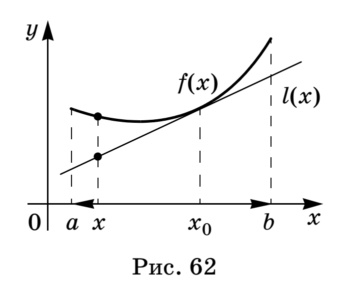


Рис. 62

является вместе с тем необходимым. Так, функция $y = x^4$ строго выпукла вниз на всей числовой прямой, однако ее вторая производная $y'' = 12x^2$ обращается в нуль при $x = 0$.

Отметим, что если функция f (строго) выпукла вверх на интервале (a, b), то функция $-f$ (строго) выпукла вниз на этом интервале и обратно, а поскольку $\frac{d^2}{dx^2}[-f(x)] = -\frac{d^2 f(x)}{dx^2}$, например, приводимое в теореме 5 достаточное условие строгой выпуклости вверх следует из содержащегося в этой же теореме достаточного условия строгой выпуклости функции вниз.

УПРАЖНЕНИЯ. 4. Доказать, что для функции

$$f(x) = \begin{cases} x \sin \frac{1}{x} & \text{при } x \neq 0, \\ 0 & \text{при } x = 0 \end{cases}$$

точка $x = 0$ не принадлежит никаким интервалам выпуклости вверх или вниз и не является концом какого-либо из этих интервалов.

5. Доказать, что функция $y = x^4$ строго выпукла вниз на всей числовой прямой.

Таким образом, выпуклость вверх или вниз функции f зависит от знака ее второй производной. Оказывается, что расположение графика дважды дифференцируемой функции относительно касательной также в определенном смысле связано со знаком второй производной.

ТЕОРЕМА 6. *Пусть функция f имеет на всем интервале (a, b) положительную (отрицательную) вторую производную:* $f''(x) > 0$ (*соответственно* $f''(x) < 0$), $x \in (a, b)$[1]. *Тогда, какова бы ни была точка $x_0 \in (a, b)$, все точки $(x, f(x))$, $x \in (a, b)$, графика функции f лежат выше (ниже) касательной, проведенной к нему в точке $(x_0, f(x_0))$, кроме, конечно, самой этой точки, которая лежит на указанной касательной*[2] (рис. 62).

[1] Отсюда следует, что функция f строго выпукла вниз (вверх) на (a, b).

[2] Если функция f, кроме того, определена и имеет одностороннюю производную в конце a или b интервала (a, b), то указанное свойство, как это видно из приводимого ниже доказательства, выполняется и для касательной в точке $(a, f(a))$, соответственно в точке $(b, f(b))$.

Действительно, уравнением касательной к графику функции в точке $(x_0, f(x_0))$ является

$$y = f'(x_0)(x - x_0) + f(x_0).$$

Обозначим правую часть этого уравнения через $L(x)$. Тогда, применив теорему Лагранжа к разности $f(x) - f(x_0)$, получим

$$f(x) - L(x) = [f(x) - f(x_0)] - f'(x_0)(x - x_0) =$$
$$= f'(\xi)(x - x_0) - f'(x_0)(x - x_0) = [f'(\xi) - f'(x_0)](x - x_0),$$

где $a < x_0 < b$, $a < x < b$, а точка ξ лежит между x и x_0.

Применив еще раз теорему Лагранжа, но уже к приращению производной, получим

$$f(x) - L(x) = f''(\eta)(\xi - x_0)(x - x_0),$$

где точка η лежит между ξ и x_0.

При $x \neq x_0$ имеем $(\xi - x_0)(x - x_0) > 0$, поскольку, как это было отмечено выше, точка ξ всегда лежит между x и x_0 и, следовательно, всегда по ту же сторону от точки x_0, что и точка x. Таким образом, знак разности $f(x) - L(x)$ совпадает, при $x \neq x_0$, со знаком $f''(\eta)$. Поэтому, если на интервале (a, b) вторая производная положительна (следовательно, она положительна и в точке η), то для всех $x \in (a, b)$, кроме точки $x = x_0$, выполняется неравенство $f(x) - L(x) > 0$; если же на интервале (a, b) вторая производная отрицательна, то для указанных точек справедливо неравенство $f(x) - L(x) < 0$. □

Поясним эту теорему исходя из несколько иных соображений. Если функция f имеет всюду на некотором интервале вторую производную, то в окрестности любой точки x_0 этого интервала можно выделить главную часть функции f в виде многочлена Тейлора второго порядка

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + o((x - x_0)^2),$$
$$x \to x_0,$$

и, следовательно, график функции f «ведет себя в окрестности точки x_0 почти как парабола»

$$y = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2,$$

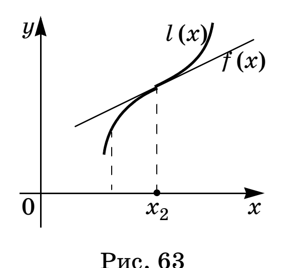


Рис. 63

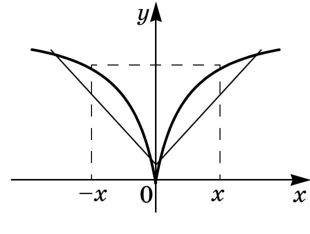


Рис. 64

которая, когда ее коэффициент при x^2, т. е. $\frac{f''(x_0)}{2}$, положителен, выпукла вниз и лежит выше любой своей касательной, в частности и выше касательной в точке $(x_0, f(x_0))$ (эта касательная к параболе является и касательной к графику функции f), а когда указанный коэффициент отрицателен, выпукла вверх и лежит ниже любой своей касательной.

Мы снова видим, как целесообразно при изучении функции в окрестности данной точки выделить с помощью формулы Тейлора главную часть функции в этой точке. В дальнейшем при решении разнообразных задач анализа мы еще неоднократно будем иметь возможность убедиться в больших возможностях и плодотворности метода выделения главной части.

Определение 6. *Пусть функция f дифференцируема при $x = x_0$ и пусть $y = L(x)$ — уравнение касательной к графику функции f в точке $(x_0, f(x_0))$. Если разность $f(x) - L(x)$ меняет знак при переходе через точку x_0, то x_0 называется точкой перегиба функции f.*

Более подробно и точно это означает, что существует такая δ-окрестность $U(x_0, \delta)$ точки x_0, что на каждом из интервалов $(x_0 - \delta, x_0)$ и $(x_0, x_0 + \delta)$ разность $f(x) - L(x)$ сохраняет постоянный знак, противоположный ее знаку на другом интервале.

Геометрический смысл точки перегиба x_0 состоит в том, что график функции f переходит в точке $(x_0, f(x_0))$ с одной стороны наклонной касательной в этой точке на другую (рис. 63).

Если x_0 — точка перегиба функции, то точка $(x_0, f(x_0))$ называется *точкой перегиба графика функции f.*

Примеры. 1. $f(x) = x^3$, $f''(x) = 6x$. Очевидно, что в этом случае $f''(x) < 0$ для $x < 0$ и $f''(x) > 0$ для $x > 0$. Поэтому на

бесконечном интервале $(-\infty, 0)$ функция $f(x) = x^3$ строго выпукла вверх, на интервале $(0, +\infty)$ она строго выпукла вниз, а точка $x = 0$ является одновременно концом интервалов выпуклости вверх и вниз. Она является и точкой перегиба, поскольку уравнением касательной в ней является $y = 0$, и при $x < 0$ имеет место неравенство $f(x) < 0$, а при $x > 0$ — неравенство $f(x) > 0$.

2. $f(x) = \sqrt[3]{x^2}$ график этой функции (рис. 64) называется *полукубической параболой*. Здесь $f''(x) = -\dfrac{2}{9\sqrt[3]{x^4}}$, поэтому для всех $x \neq 0$ справедливо неравенство $f''(x) < 0$. Следовательно, интервалы $(-\infty, 0)$ и $(0, +\infty)$ являются промежутками строгой выпуклости вверх. Вместе с тем при любом $x \neq 0$ имеем

$$f(x) > 0 = f(0) \text{ и } f(x) = f(-x).$$

Это означает, что точка $(0, 0)$ графика функции f лежит под любой его хордой с концами в точках $(-x, f(-x))$ и $(x, f(x))$. Поэтому точка $x = 0$ не принадлежит никакому интервалу выпуклости вверх (интервалов выпуклости вниз у этой функции нет).

График функции $f(x) = \sqrt[3]{x^2}$ в точке $(0, 0)$ имеет вертикальную касательную, и его ветви, для которых $x > 0$ и $x < 0$, лежат по разные стороны от нее. Однако $x = 0$ не является точкой перегиба, поскольку, в силу вертикальности касательной в этой точке, функция f не дифференцируема в ней, следовательно, $x = 0$ не удовлетворяет условиям определения 6.

Образно говоря, график полукубической параболы «не перегибается» при переходе через касательную в точке $(0, 0)$, а «возвращается назад»; поэтому точки такого типа называются *точками возврата*.

ТЕОРЕМА 7 (необходимое условие, выполняющееся в точке перегиба). *Если в точке перегиба функции существует вторая производная, то она равна нулю.*

Доказательство. Действительно, пусть функция f имеет в точке x_0 вторую производную и, как и выше, $y = L(x)$ — уравнение касательной к графику функции f в точке $(x_0, f(x_0))$, т. е.

$$L(x) \equiv f(x_0) + f'(x_0)(x - x_0).$$

Тогда, в силу формулы Тейлора,

$$f''(x) - L(x) = \frac{f''(x_0)}{2}(x - x_0)^2 + o((x - x_0)^2),\ x \to x_0.$$

Если $f''(x_0) \neq 0$, то знак разности $f(x) - L(x)$ в некоторой окрестности точки x_0 совпадает со знаком числа $f''(x_0)$. В этом случае разность $f(x) - L(x)$ не меняет знака в точке x_0 и, следовательно, эта точка не является точкой перегиба. Итак, если x_0 — точка перегиба функции f, то $f''(x_0) = 0$. □

З а м е ч а н и е 2. Подобно тому, как все точки экстремума функции принадлежат множеству точек, в которых производная либо равна нулю, либо не существует, так и все точки перегиба функции (дважды дифференцируемой для всех значений аргумента, кроме, быть может, конечного числа его значений) входят во множество точек, в которых вторая производная либо равна нулю, либо не существует.

ТЕОРЕМА 8 (первое достаточное условие для точки перегиба). *Если функция f, дифференцируемая в точке x_0, дважды дифференцируема в некоторой проколотой окрестности $\mathring{U}(x, \delta)$ этой точки и вторая производная f'' функции f меняет знак при переходе аргумента через значение x_0 (т. е. либо $f''(x) < 0$ при $x_0 - \delta < x < x_0$ и $f''(x) > 0$ при $x_0 < x < x_0 + \delta$, либо $f''(x) > 0$ при $x_0 - \delta < x < x_0$ и $f''(x) < 0$ при $x_0 < x < x_0 + \delta$), то x_0 является точкой перегиба функции f.*

В самом деле, запишем, как и выше, уравнение касательной $y = f'(x_0)(x - x_0) + f(x_0)$ в виде $y = L(x)$. При доказательстве теоремы 6 было показано, что

$$f(x) - L(x) = f''(\eta)(\xi - x_0)(x - x_0),$$

где точки x и ξ лежат по одну сторону от точки x_0, поэтому при $x \neq x_0$ имеем $(\xi - x_0)(x - x_0) > 0$ и, следовательно,

$$\operatorname{sign}\,[f(x) - L(x)] = \operatorname{sign} f''(\eta).$$

Точка η лежит между ξ и x_0, т. е. по ту же сторону от x_0, что и точка x. Отсюда имеем, что если f'' меняет знак при переходе аргумента через точку x_0, то разность $f(x) - L(x)$ меняет знак и, следовательно, x_0 является точкой перегиба. □

ТЕОРЕМА 9 (второе достаточное условие для точки перегиба). *Пусть* $f''(x_0) = 0$, *а* $f'''(x_0) \neq 0$; *тогда точка* x_0 *является точкой перегиба функции* f.

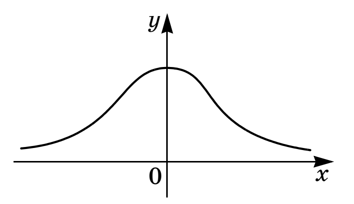
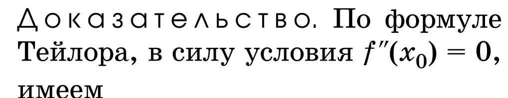


Рис. 65



Доказательство. По формуле Тейлора, в силу условия $f''(x_0) = 0$, имеем

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f'''(x_0)}{3!}(x - x_0)^3 + o((x - x_0)^3),$$

$x \to x_0$, и, поскольку $L(x) = f(x_0) + f'(x_0)(x - x_0)$, то

$$f(x) - L(x) = \frac{f'''(x_0)}{3!}(x - x_0)^3 + o((x - x_0)^3),\ x \to x_0.$$

Отсюда следует (см. замечание о бесконечно малых перед доказательством теоремы 4 этого параграфа), что знак разности $f(x) - L(x)$ меняется при изменении знака $x - x_0$. Это и означает, что x_0 является точкой перегиба. □

Пример 3. Рассмотрим функцию $f(x) = e^{-x^2}$ и найдем ее точки перегиба. Имеем

$$f'(x) = -2xe^{-x^2},$$

$$f''(x) = -2e^{-x^2} + 4x^2e^{-x^2} = 4\left(x^2 - \frac{1}{2}\right)e^{-x^2} =$$

$$= 4\left(x + \frac{1}{\sqrt{2}}\right)\left(x - \frac{1}{\sqrt{2}}\right)e^{-x^2}.$$

Отсюда видно, что вторая производная функции f обращается в нуль в точках $x = \pm 1/\sqrt{2}$ и при переходе через них меняет знак. Следовательно, согласно теореме 8, эти точки являются точками перегиба функции f (рис. 65).

Задача 11. Доказать, что если функция f непрерывна на интервале (a, b) и если для любых точек x_1 и x_2, $a < x_1 < x_2 < b$, выполняется неравенство

$$\frac{f(x_1) + f(x_2)}{2} \leqslant f\left(\frac{x_1 + x_2}{2}\right),$$

то (a, b) является интервалом выпуклости вверх для функции f.

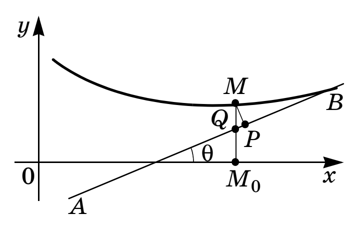


Рис. 66

Задача 12. Доказать следующие ниже утверждения. Для того чтобы дифференцируемая функция была выпуклой вверх (вниз) на некотором интервале, необходимо и достаточно, чтобы ее производная убывала (возрастала) на нем. Для того чтобы дифференцируемая функция была строго выпуклой вверх (вниз) на некотором интервале, достаточно, чтобы ее производная строго убывала (строго возрастала) на нем.

14.4. Асимптоты

Определение 7. *Пусть функция* $f(x)$ *определена для всех* $x > a$ (*соответственно для всех* $x < a$). *Если существуют такие числа* k *и* l, *что* $\lim\limits_{x \to +\infty} [f(x) - (kx + l)] = 0$ (*при* $x \to -\infty$), *то прямая*

$$y = kx + l \tag{14.9}$$

называется асимптотой графика функции $f(x)$ *при* $x \to +\infty$ (*при* $x \to -\infty$).

Иногда вместо «асимптота графика функции» говорят короче: «асимптота функции».

Существование асимптоты графика функции означает, что при $x \to +\infty$ (или $x \to -\infty$) функция ведет себя «почти как линейная функция», т. е. отличается от линейной функции на бесконечно малую.

Найдем, например, асимптоту графика функции $y = \frac{x^2 - 3x - 2}{x + 1}$. Разделив числитель на знаменатель по правилу деления многочленов, получим $y = x - 4 + \frac{2}{x+1}$. Так как $\frac{2}{x+1} = o(1)$ при $x \to \pm\infty$, то прямая $y = x - 4$ является асимптотой графика данной функции как при $x \to +\infty$, так и при $x \to \infty$.

Рассмотрим геометрический смысл асимптоты. Пусть $M = (x, f(x))$ — точка графика функции f, M_0 — проекция этой точки на ось Ox, AB — асимптота (14.9), θ — угол между асимптотой и положительным направлением оси Ox, $\theta \neq \pi/2$, MP — перпендикуляр, опущенный из точки M на асимптоту AB, Q — точка пересечения прямой MM_0 с асимптотой AB (рис. 66). Тогда $MM_0 = f(x)$, $QM_0 = kx + l$, $MQ = MM_0 - QM_0 =$

$= f(x) - (kx + l)$, $MP = MQ\cos\theta$. Таким образом, MP отличается от MQ лишь на не равный нулю множитель $\cos\theta$, поэтому условия $MQ \to 0$ и $MP \to 0$ при $x \to +\infty$ (соответственно при $x \to -\infty$) эквивалентны, т. е. если $\lim\limits_{x \to +\infty} MQ = 0$, то и $\lim\limits_{x \to +\infty} MP = 0$, и наоборот. Отсюда следует, что асимптота может быть определена как прямая, расстояние до которой от графика функции, т. е. отрезок MP, стремится к нулю, когда точка $M = (x, f(x))$ «стремится, оставаясь на графике, в бесконечность» (при $x \to +\infty$ или, соответственно, при $x \to -\infty$).

Укажем теперь общий метод отыскания асимптоты (14.9), т. е. способ определения коэффициентов k и l в уравнении (14.9). Будем рассматривать для определенности лишь случай $x \to +\infty$ (при $x \to -\infty$ рассуждения проводятся аналогично). Пусть график функции f имеет асимптоту (14.9) при $x \to +\infty$. Тогда, по определению,

$$f(x) = kx + l + o(1). \tag{14.10}$$

Разделим обе части равенства (14.10) на x и перейдем к пределу при $x \to +\infty$. Тогда

$$\lim_{x \to +\infty} \frac{f(x)}{x} = k. \tag{14.11}$$

Используя найденное значение k, получим из (14.10) для определения l формулу

$$l = \lim_{x \to +\infty} (f(x) - kx). \tag{14.12}$$

Справедливо и обратное утверждение: если существуют такие числа k и l, что выполняется условие (14.12), то прямая $y = kx + l$ является асимптотой графика функции $f(x)$. В самом деле, из (14.12) имеем

$$\lim_{x \to +\infty} [f(x) - (kx + l)] = 0,$$

т. е. прямая $y = kx + l$ действительно удовлетворяет определению асимптоты, иначе говоря, выполняется условие (14.10).

Таким образом, формулы (14.11) и (14.12) сводят задачу отыскания асимптот (14.9) к вычислению пределов определенного вида. Более того, мы показали, что если существует представление функции f в виде (14.10), то k и l выражаются по формулам (14.11) и (14.12). Следовательно, если существует представление (14.10), то оно единственно.

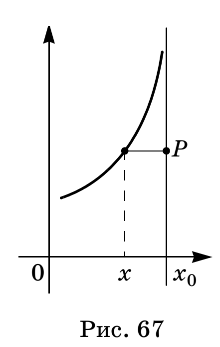


Рис. 67

Найдем по этому правилу асимптоту графика функции $f(x) = \dfrac{x^2 - 3x - 2}{x + 1}$, найденную выше другим способом:

$$k = \lim_{x \to \infty} \frac{f(x)}{x} = \lim_{x \to \infty} \frac{x^2 - 3x - 2}{x(x + 1)} = 1,$$

$$l = \lim_{x \to \infty} \left(\frac{x^2 - 3x - 2}{x + 1} - x \right) = \lim_{x \to \infty} \frac{-4x - 2}{x + 1} = -4,$$

т. е. мы, как и следовало ожидать, получили то же уравнение асимптоты $y = x - 4$, как при $x \to +\infty$, так и при $x \to -\infty$.

В виде (14.9) может быть записано уравнение любой прямой, непараллельной оси Oy. Естественно распространить определение асимптоты и на прямые, параллельные оси Oy.

Определение 8. *Пусть функция f определена в некоторой окрестности или проколотой окрестности точки x_0* (*быть может, односторонней*) *и пусть выполнено хотя бы одно из условий*

$$\lim_{x \to x_0 - 0} f(x) = \infty, \text{ или } \lim_{x \to x_0 + 0} f(x) = \infty. \qquad (14.13)$$

Тогда прямая $x = x_0$ (рис. 67) *называется вертикальной асимптотой графика функции f* (*в отличие от асимптоты вида* (14.9), *которая называется также наклонной асимптотой*).

В случае вертикальной асимптоты, как и в случае наклонной, расстояние $MP = x - x_0$ между точкой M и прямой $x = x_0$ стремится к нулю, если точка $M(x, f(x))$ стремится вдоль графика в бесконечность, т. е. когда $x \to x_0 - 0$ или $x \to x_0 + 0$.

Чтобы найти вертикальные асимптоты графика функции f, надо найти такие значения x_0, для которых выполняется одно или оба условия (14.13). Например, функция $y = \dfrac{x^2 - 3x - 2}{x + 1}$ имеет вертикальную асимптоту $x = -1$. Вообще если $f(x) = P(x)/Q(x)$ — рациональная функция (($P(x)$ и $Q(x)$ — многочлены), $Q(x_0) = 0$, $P(x_0) \neq 0$, то прямая $x = x_0$ является асимптотой графика функции $f(x)$.

14.5. Построение графиков функций

Изучение заданной функции и построение ее графика с помощью развитого нами аналитического аппарата целесообразно проводить в следующем порядке.

1. Определить область существования функции, область непрерывности и точки разрыва.

2. Найти асимптоты.

3. Приблизительно, вчерне, нарисовать график функции.

4. Вычислить первую, а если нужно, и вторую производную (без производных более высокого порядка часто удается обойтись).

5. Найти точки, в которых первая и вторая производные либо не существуют, либо равны нулю.

6. Составить таблицу изменения знака первой и второй производных. Определить промежутки возрастания, убывания, выпуклости вверх или вниз функции, найти точки экстремума (в том числе концевые) и точки перегиба.

7. Окончательно вычертить график.

При этом чем бóльшую точность графика мы хотим достигнуть, тем больше, вообще говоря, необходимо найти точек, лежащих на нем. Обычно целесообразно найти (быть может, с определенной точностью) точки пересечения графика с осями координат и точки, соответствующие экстремумам функции; другие точки находятся по мере потребности.

В случае очень громоздких выражений для второй производной иногда приходится ограничиваться рассмотрением тех свойств графика, которые можно изучать лишь с помощью первой производной.

Пример 1. Построим график функции $f(x) = \frac{x^2 - 3x - 2}{x + 1}$.

Эта функция определена и непрерывна для всех $x \neq -1$. Она, как известно (см. п. 14.4), имеет асимптоты $y = x - 4$ и $x = -1$, причем $\lim\limits_{x \to -1 - 0} f(x) = -\infty$, $\lim\limits_{x \to -1 + 0} f(x) = +\infty$. Было отмечено, что $f(x) = x - 4 + \frac{2}{x + 1}$, поэтому $f(x) > x - 4$ при $x > -1$ (график функции находится выше асимптоты) и $f(x) < x - 4$ при $x < -1$ (график лежит ниже асимптоты).

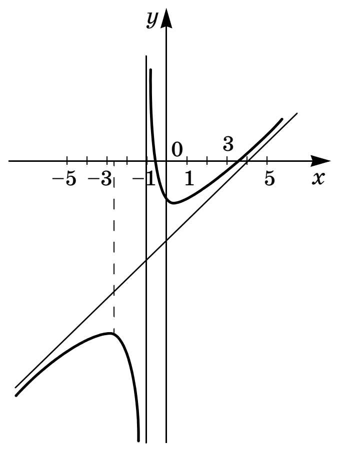


Рис. 68

График функции $f(x)$ пересекает ось Ox в точках, в которых $x^2 - 3x - 2 = 0$, т. е. при $x_1, x_2 = (3 \pm \sqrt{17})/2$ или приблизительно в точках $x_1 \approx 3{,}5$, $x_2 \approx -0{,}5$. Ось Oy график пересекает в точке $y = -2$. Это позволяет нарисовать график функции $f(x)$ в виде, указанном на рисунке 68.

Дальнейшее исследование имеет своей целью нахождение экстремумов точек перегиба и интервалов, выпуклости вверх или вниз графика функции. Для этого найдем производные y' и y'':

$$y' = \frac{x^2 + 2x - 1}{(x+1)^2}, \quad y'' = \frac{4}{(x+1)^3}.$$

Отсюда видно, что $y' = 0$ при $x = -1 - \sqrt{2} \approx -2{,}4$ и $x = -1 + \sqrt{2} \approx 0{,}4$. В точке $x = -1$ производные y' и y'' не существуют (даже функция f в этой точке не определена).

Составим таблицу изменения знака первой и второй производных в зависимости от изменения аргумента, включив в нее критические точки:

x		$-1-\sqrt{2}$		-1		$-1+\sqrt{2}$	
y'	+	0	−	Не существует	−	0	+
y''	−	−	−	Не существует	+	+	+

Из этой таблицы видно, что функция $f(x)$ имеет в точке $x = -1 + \sqrt{2}$ строгий минимум, а в точке $x = -1 - \sqrt{2}$ — строгий максимум; при $x < -1$ функция строго выпукла вверх, а при $x > -1$ — строго выпукла вниз. Точек перегиба нет, так как при $x = -1$ функция разрывна.

В дальнейшем для краткости подобные таблицы будем называть *таблицами поведения функций* и иногда сразу

отмечать в них точки экстремума, точки перегиба и интервалы выпуклости.

Мы нашли общий характер поведения функции. Чтобы построить график более точно, надо найти ряд точек графика, как это отмечалось выше.

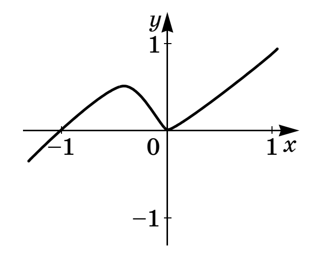


Рис. 69

Пример 2. Построим график функции

$$f(x) = (x+1)^3 \sqrt[3]{x^2}.$$

Область определения этой функции — множество всех действительных чисел, причем она непрерывна в каждой точке и поэтому не имеет вертикальных асимптот. Из того, что

$$\lim_{x \to \pm\infty} \frac{f(x)}{x} = \infty,$$

следует, что нет и наклонных асимптот.

Для построения графика вчерне заметим, что:

1) $f(x)$ обращается в нуль в точках $x = -1$ и $x = 0$;

2) $f > 0$ при $x > -1$, $x \neq 0$;

3) $f < 0$ при $x < -1$;

4) $\lim\limits_{x \to +\infty} f(x) = +\infty$ и $\lim\limits_{x \to -\infty} f(x) = -\infty$.

Приблизительный вид графика функции, который можно нарисовать на основании этих замечаний, изображен на рисунке 69.

Проведем теперь более подробное исследование функции с помощью производных. Найдем y' и y'':

$$y' = \frac{(x+1)^2(11x+2)}{3\sqrt[3]{x}},$$

$$y'' = \frac{2(x+1)(44x^2+16x-1)}{9x\sqrt[3]{x}}.$$

Отсюда видно, что $y' = 0$ при $x = -1$ и $x = -2/11$; $y'' = 0$ при $x = -1$, а также когда $44x^2 + 16x - 1 = 0$, т. е. приблизительно при $x_1 \approx -9/22$ и $x_2 \approx 1/22$. При $x = 0$ производные y' и y'' не существуют.

Составляем таблицу поведения функции:

<table>
<tr><th>x</th><th>y'</th><th>y''</th><th>Интервалы монотонности и точки экстремума</th><th>Интервалы выпуклости и точки перегиба</th></tr>
<tr><td>$(-\infty, -1)$</td><td>+</td><td>−</td><td rowspan="5">Возрастание</td><td>Выпуклость вверх</td></tr>
<tr><td>-1</td><td>0</td><td>0</td><td>Точка перегиба</td></tr>
<tr><td>$(-1, x_1)$</td><td>+</td><td>+</td><td>Выпуклость вниз</td></tr>
<tr><td>x_1</td><td>+</td><td>0</td><td>Точка перегиба</td></tr>
<tr><td>$\left(x_1, -\frac{2}{11}\right)$</td><td>+</td><td>−</td><td rowspan="5">Выпуклость вверх</td></tr>
<tr><td>$-\frac{2}{11}$</td><td>0</td><td>−</td><td>Максимум</td></tr>
<tr><td>$\left(-\frac{2}{11}, 0\right)$</td><td>−</td><td>−</td><td>Убывание</td></tr>
<tr><td>0</td><td>Не существует</td><td>Не существует</td><td>Минимум</td></tr>
<tr><td>$(0, x_2)$</td><td>+</td><td>−</td><td rowspan="3">Возрастание</td></tr>
<tr><td>x_2</td><td>+</td><td>0</td><td>Точка перегиба</td></tr>
<tr><td>$(x_2, +\infty)$</td><td>+</td><td>+</td><td>Выпуклость вниз</td></tr>
</table>

Теперь график функции $y = (x + 1)\sqrt[3]{x}$ можно нарисовать более точно. Его вид изображен на рисунке 70. Как видно, исследование с помощью производных позволило существенно уточнить вид графика (ср. рис. 69 и 70).

Развитый аппарат позволяет строить и графики функций, локально заданных параметрически: $x = x(t)$, $y = y(t)$. Здесь не предполагается, что пара функций $x = x(t)$, $y = y(t)$ определяет однозначно одну функцию вида $y = y(x)$ или $x = x(t)$. Под графиком параметрически заданной функции подразумевается объединение графиков всех функций вида $y = f(x)$ и $x = g(x)$, задаваемых формулами $x = x(t)$, $y = y(t)$.

Сделаем несколько предварительных замечаний. Для нахождения асимптот, параллельных оси Oy, необходимо найти такие значения t_0[1], для которых существует конечный

[1] Здесь и далее t_0 — число или одна из бесконечностей $+\infty$, $-\infty$.

предел $\lim\limits_{t\to t_0+0} x(t) = a$ или $\lim\limits_{t\to t_0-0} x(t) = a$,

а $\lim\limits_{t\to t_0+0} y(t)$, соответственно $\lim\limits_{t\to t_0-0} y(t)$,

равен $+\infty$ или $-\infty$. Если такие значения t_0 существуют, то

$$x = a \qquad (14.14)$$

является уравнением искомой асимптоты.

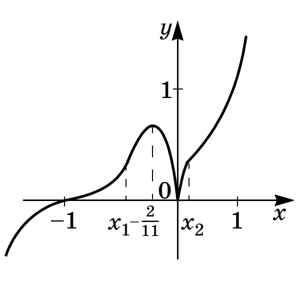


Рис. 70

Аналогично, нахождение асимптот, параллельных оси Ox, сводится к определению таких значений t_0, для которых существует конечный предел

$\lim\limits_{t\to t_0+0} y(t) = b$ или $\lim\limits_{t\to t_0-0} y(t) = b$, а $\lim\limits_{t\to t_0+0} x(t)$, соответственно $\lim\limits_{t\to t_0-0} x(t)$, равен $+\infty$ или $-\infty$. Если окажется, что такие значения t_0 существуют, то

$$y = b \qquad (14.15)$$

является уравнением искомой асимптоты.

Наконец, для нахождения асимптот, не параллельных ни оси Ox, ни оси Oy, надо найти такие значения t_0, для которых пределы $\lim\limits_{t\to t_0+0} x(t)$ и $\lim\limits_{t\to t_0+0} y(t)$ (или $\lim\limits_{t\to t_0-0} x(t)$ и $\lim\limits_{t\to t_0-0} (y(t))$ равны $+\infty$ или $-\infty$ и существует конечный предел $\lim\limits_{t\to t_0+0} \frac{y(t)}{x(t)} = k - 0$ $\left(\text{соответственно } \lim\limits_{t\to t_0-0} \frac{y(t)}{x(t)} = k\right)$. Если для этого значения t_0, кроме того, существует конечный предел $\lim\limits_{t\to t_0+0} [y(t) - kx(t)] = l$ (соответственно $\lim\limits_{t\to t_0-0} [y(t) - kx(t)] = l$), то прямая

$$y = kx + l \qquad (14.16)$$

является асимптотой графика рассматриваемой функции.

УПРАЖНЕНИЕ 6. Вывести уравнение асимптот (14.14), (14.15) и (14.16), исходя из того, что асимптотой называется прямая, расстояние до которой от точки $(x(t), y(t))$ графика функции, заданной параметриче-

ски: $x = x(t)$, $y = y(t)$, стремится к нулю, когда точка стремится, оставаясь на графике функции, в бесконечность, т. е. когда $\sqrt{x^2(t) + y^2(t)} \to \infty$ при $t \to t_0 + 0$ или $t \to t_0 - 0$.

При предварительном построении графика функции, заданной параметрически, часто бывает полезно построить сначала отдельно графики функций $x = x(t)$ и $y = y(t)$.

Для определения промежутков возрастания и убывания функции, заданной параметрически, нахождения ее экстремумов, точек перегиба, а также интервалов выпуклости вверх и вниз надо использовать выражения для производных y'_x и y''_{xx} через производные x'_t, y'_t, y''_{tt} и x''_{tt}. При этом следует иметь в виду, что уравнения $x = x(t)$, $y = y(t)$, вообще говоря, не определяют однозначно функцию вида $y = y(x)$, так что при исследовании графика функции надо все время внимательно следить за тем, какая «ветвь» графика рассматривается. Иногда, наоборот, полезнее рассматривать x как функцию от y.

Пример 3. Построим график функции

$$x = \frac{t^2 + 1}{4(1 - t)}, \quad y = \frac{t}{1 + t}. \qquad (14.17)$$

Параметрическое представление имеет смысл для всех t, кроме $t = \pm 1$. Асимптоты, параллельные оси Ox, получаются при $t = 1$ и $t = \pm\infty$; их уравнения соответственно $y = 1/2$ и $y = 1$. Асимптота, параллельная оси Oy, получается при $t = -1$; ее уравнение $x = 1/4$. Наклонных асимптот в данном случае нет.

Для построения графика вчерне полезно составить таблицу изменения знаков переменных x и y в зависимости от изменения t; в нее могут быть включены и некоторые характерные значения x и y. Так, в данном случае полезна следующая таблица:

t	$-\infty$		-1		0		1		$+\infty$
x	$+\infty$	$+$	$1/4$	$+$	$1/4$	$+$	∞	$-$	$-\infty$
y	1	$+$	∞	$-$	0	$+$	$1/2$	$+$	1

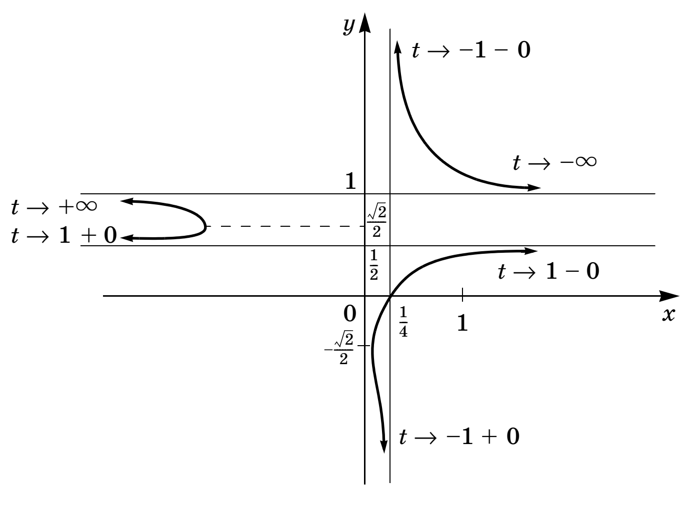


Рис. 71

Теперь строим график (рис. 71). Для наглядности на графике указано, как ветви графика соответствуют изменению параметра.

Далее имеем

$$x'_t = \frac{1 + 2t - t^2}{4(1 - t)^3}, \quad y'_t = \frac{1}{(1 + t)^2},$$

поэтому

$$x'_y = \frac{(1 + t)^2(1 + 2t - t^2)}{4(1 - t^2)}. \qquad \textbf{(14.18)}$$

В данном случае лучше рассматривать x как функцию от y, а не наоборот, так как из построенного графика видно, что естественно ожидать, что x определяется однозначно как функция y, $y \neq 1/2$ и $y \neq 1$. Это легко установить и аналитически, если заметить, что уравнение $y = \frac{t}{1 + t}$ однозначно разрешимо относительно параметра t.

Из (14.18) видно, что $x'_y = 0$ при $t = -1$ и когда $1 + 2t - t^2 = 0$, т. е. при $t = 1 + \sqrt{2}$ и $t = 1 - \sqrt{2}$. Значению $t = -1$ не соответствует никакая точка графика, а при $t = 1 + \sqrt{2}$ и $t = 1 - \sqrt{2}$ имеем соответственно

$$y = \frac{1 + \sqrt{2}}{2 + \sqrt{2}} = \frac{\sqrt{2}}{2} \quad \text{и} \quad y = \frac{1 - \sqrt{2}}{2 - \sqrt{2}} = -\frac{\sqrt{2}}{2}.$$

Составим теперь таблицу изменения знака производной x'_y; эта таблица позволяет найти и точки экстремума:

t	$-\infty$		-1		$1-\sqrt{2}$		1		$1+\sqrt{2}$		$+\infty$
y	1		∞		$-\frac{\sqrt{2}}{2}$		$\frac{1}{2}$		$\frac{\sqrt{2}}{2}$		1
x'_y		$-$	0	$-$	0	$+$		$+$	0	$-$	
Экстремумы					Минимум				Максимум		

Из таблицы видно, что в точке $y = \sqrt{2}/2$ функция $x = x(y)$ имеет максимум, в точке $y = -\sqrt{2}/2$ — минимум и строго монотонна на следующих интервалах: $(-\infty, -\sqrt{2}/2)$, $(-\sqrt{2}/2, 1/2)$, $(1/2, \sqrt{2}/2)$, $(\sqrt{2}/2, 1)$, $(1, +\infty)$.

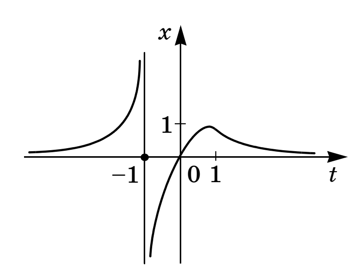
Необходимо обратить внимание на то, что, взяв y за независимую переменную, x — за зависимую, т. е. взяв ось Oy за первую координатную ось, а Ox — за вторую, мы получим систему координат, ориентированную противоположно рассматриваемой все время системе координат, у которой первой осью является Ox, а второй — Oy. Читателю полезно убедиться, что доказанные выше критерии, например, для наличия экстремумов и точек перегиба геометрически не связаны с той или иной ориентацией осей координат.

Для исследования выпуклости и точек перегиба функции $x(y)$ найдем x''_{yy}:

$$x''_{yy} = (x'_y)'_t t'_y = \frac{(1+t)^3(3+3t-3t^2+t^3)}{2(1-t)^3}.$$

Производная x''_{yy} равна нулю при $t = -1$ и для тех t, для которых $P(t) \equiv 3 + 3t - 3t^2 + t^3 = 0$.

В силу того, что $P'(t) = 3(t-1)^2 \geqslant 0$, причем $P' = 0$ только в одной точке $t = 1$, функция $P(t)$ строго монотонно возрастает на всей вещественной оси (почему?). Следовательно, существует единственное t_0 такое, что $P(t_0) = 0$. При этом $P(0) = 3 > 0$, а $P(-1) = -4 < 0$, откуда $-1 < t_0 < 0$. Если $y_0 = \dfrac{t_0}{1+t_0}$, то, очевидно, $-\infty < y_0 < 0$ (можно, конечно, получить и более точную оценку для y_0, выбирая более близкие t_1 и t_2 такие, что $P(t_1) < 0$,



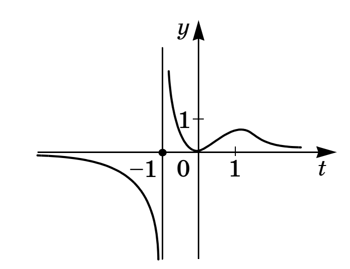


Рис. 72

$P(t_2) > 0$). Составим теперь таблицу изменения второй производной x''_{yy} и определим с ее помощью интервалы выпуклости вверх и вниз, а также точки перегиба (см. табл. на с. 384).

График функции (14.17) исследован. Он изображен на рисунке 71.

Пример 4. Построим график функции

$$x = \frac{t}{1 + t^3}, \quad y = \frac{t^2}{1 + t^3}. \qquad (14.19)$$

Асимптот, параллельных осям координат, в данном случае нет; так как $x \to \infty$ и $y \to \infty$ при $t \to -1$, то, возможно, существует наклонная асимптота. Для ее нахождения вычислим соответствующие пределы:

$$\lim_{t \to -1} \frac{y}{x} = \lim_{t \to -1} t = -1, \text{ т. е. } k = -1,$$

$$\lim_{t \to -1} (y - kx) = \lim_{t \to -1} \left(\frac{t^2}{1 + t^2} + \frac{t}{1 + t^3}\right) = \lim_{t \to -1} \frac{t}{t^2 - t + 1} = -\frac{1}{3}.$$

Отсюда следует, что наклонная асимптота существует и что ее уравнением является $y = -x - \frac{1}{3}$.

Построим приблизительно графики функций $x(t)$ и $y(t)$; для этого предварительно найдем производные:

$$x'_t = \frac{1 - 2t^3}{(1 + t^3)^3}, \quad y'_t = \frac{t(2 - t^3)}{(1 + t^3)^2}. \qquad (14.20)$$

Производная x'_t обращается в нуль при $t = 1/\sqrt[3]{2}$, при этом знак производной изменяется с «+» на «−», поэтому это точка максимума; производная y'_t обращается в нуль при $t = 0$, меняя знак с «−» на «+» (значит, это точка минимума), и при $t = 1/\sqrt[3]{2}$, меняя знак с «+» на «−» (следовательно, это также точка максимума). Из этих замечаний следует, что графики функций $x(t)$ и $y(t)$ имеют вид, изображенный на рисунке 72.

t	y	x''_y	Интервалы выпуклости	Точки перегиба и точки разрыва
$-\infty$	1			Точка разрыва
$(-\infty, -1)$	$(1, +\infty)$	+	Выпуклость вниз	
-1	∞			
$(-1, t_0)$	$(-\infty, y_0)$	–	Выпуклость вверх	
t_0	y_0	0		Точка перегиба
$(t_0, 1)$	$\left(y_0, \frac{1}{2}\right)$	+	Выпуклость вниз	
1	$\frac{1}{2}$	Не существует		Точка разрыва
$(1, +\infty)$	$\left(\frac{1}{2}, 1\right)$	–	Выпуклость вверх	
$+\infty$	1			Точка разрыва

По этим графикам, зная уравнение асимптоты, можно найти приблизительно график искомой функции (14.19). Он имеет вид, изображенный на рисунке 73.

Исследование производной y'_x позволит уточнить размеры «петли», образуемой графиком. Из (14.20) имеем $y'_x = \frac{t(2-t^3)}{1-2t^3}$. Отсюда следует:

1) $y'_x = 0$ при $t = 0$ и $t = \sqrt[3]{2}$, т. е. касательная к графику параллельна оси Ox в точках $(0, 0)$ и $(\sqrt[3]{2}/3, \sqrt[3]{4}/3)$;

2) $y'_x = \infty$ при $t = 1/\sqrt[3]{2}$ и $t = \infty$, т. е. касательная параллельна оси Oy в точках $(\sqrt[3]{4}/3, \sqrt[3]{2}/3)$ и $(0, 0)$. Таким образом, точке $(0, 0)$ (являющейся, как говорят, точкой самопересечения графика) соответствуют два значения параметра

$t = 0$ и $t = \infty$, если только доопределить функции (14.19), положив $x(\infty) = 0$, $y(\infty) = 0$. В этой точке две части графика имеют соответственно своими касательными координатные оси.

График функции (14.19) называется *декартовым*[1] *листом*. Из формул (14.19) нетрудно получить его неявное задание: $x^3 + y^3 - xy = 0$.

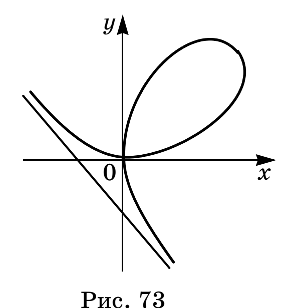


Рис. 73

§ 15.

Векторная функция

15.1. Понятие предела и непрерывности для векторной функции

В этом параграфе рассматриваются функции, значениями которых являются векторы, а аргументами — числа. Такие функции называются *векторными функциями* (числового аргумента). Они обозначаются в тексте полужирным шрифтом: $\boldsymbol{r}(t)$, $t \in X$, где X — некоторое числовое множество.

В этом определении в зависимости от рассматриваемых задач под значениями $\boldsymbol{r}$ векторных функций могут пониматься как свободные векторы, так и векторы с закрепленными началами. Если начала всех векторов закреплены в одной и той же точке, то такие векторы называются *радиусами-векторами*.

Если в пространстве задана прямоугольная система координат, то, как хорошо известно, каждому вектору соответствует упорядоченная тройка действительных чисел — его координат и, наоборот, каждой упорядоченной тройке чисел соответствует вектор, для которого числа, входящие в эту тройку, являются его координатами. Поэтому задание векторной функции эквивалентно заданию трех скалярных (числовых) функций $x(t)$, $y(t)$, $z(t)$, являющихся его координатами: $\boldsymbol{r}(t) = (x(t), y(t), z(t))$, $t \in X$.

---

[1] Р. Декарт (1596—1650) — французский философ, математик, физик, физиолог.

Если при всех $t \in X$ имеем $z(t) = 0$, то векторная функция задается двумя координатами $x(t)$ и $y(t)$; в этом случае пишут $\boldsymbol{r}(t) = (x(t), y(t))$.

Длина всякого вектора $\boldsymbol{\rho}$ обозначается через $|\boldsymbol{\rho}|$. Будем предполагать известными основные алгебраические свойства векторов, понятие скалярного и векторного произведений, а также свойства этих произведений. Скалярное произведение векторов $\boldsymbol{a}$ и $\boldsymbol{b}$ обозначается через $\boldsymbol{ab}$ или $(\boldsymbol{a}, \boldsymbol{b})$, а векторное — через $\boldsymbol{a} \times \boldsymbol{b}$ или $[\boldsymbol{a}, \boldsymbol{b}]$.

Введем понятие предела, непрерывности, производной и дифференциала для векторных функций.

Определение 1. *Вектор $\boldsymbol{a}$ называется пределом векторной функции $\boldsymbol{r}(t)$, $t \in X$, при $t \to t_0$ (или в точке $t = t_0$), если*

$$\lim_{t \to t_0} |\boldsymbol{r}(t) - \boldsymbol{a}| = 0. \tag{15.1}$$

В этом случае пишут

$$\lim_{t \to t_0} \boldsymbol{r}(t) = \boldsymbol{a}. \tag{15.2}$$

В этом определении $|\boldsymbol{r}(t) - \boldsymbol{a}|$ — числовая функция. Таким образом, понятие предела векторной функции сводится к понятию предела скалярной функции (15.1). Вспомнив определение этого понятия, получим, что (15.2) означает, что для любого $\varepsilon > 0$ существует такое $\delta > 0$, что для всех

$$t \in X \cap U(t_0; \delta) \tag{15.3}$$

выполняется неравенство

$$|\boldsymbol{r}(t) - \boldsymbol{a}| < \varepsilon. \tag{15.4}$$

Как и в случае скалярных функций, будем предполагать, что t_0 — точка прикосновения (конечная или бесконечно удаленная) множества X. Если t_0 — конечная точка, то условие (15.3) можно записать в виде

$$|t - t_0| < \delta, \ \ t \in X,$$

а если t_0 — одна из бесконечно удаленных точек ∞, $+\infty$ или $-\infty$, то соответственно в одном из следующих трех видов:

$$|t| > \frac{1}{\delta}, \ \ t > \frac{1}{\delta} \ \text{ или } \ t < -\frac{1}{\delta},$$

где всюду $t \in X$.

Если начало всех векторов $\boldsymbol{r}(t)$ поместить в одну точку (например, начало координат), то условие (15.4) будет означать, что концы всех векторов $\boldsymbol{r}(t)$ при $t \in X \cap U(t_0; \delta)$ (см. 15.3)) лежат в шаре радиуса ε с центром в конце вектора $\boldsymbol{a}$ (рис. 74).

Рис. 74

Если $\boldsymbol{r}(t) = (x(t), y(t), z(t))$ и $\boldsymbol{a} = (a_1, a_2, a_3)$, то

$$|\boldsymbol{r}(t) - \boldsymbol{a}| = \sqrt{[x(t) - a_1]^2 + [y(t) - a_2]^2 + [z(t) - a_3]^2} \qquad (15.5)$$

и, следовательно,

$$|x(t) - \boldsymbol{a}_1| \leqslant |\boldsymbol{r}(t) - a|, |y(t) - \boldsymbol{a}_2| \leqslant \leqslant |\boldsymbol{r}(t) - a|, |z(t) - \boldsymbol{a}_3| \leqslant |\boldsymbol{r}(t) - a|. \qquad (15.6)$$

Поэтому

$$\lim_{t \to t_0} \boldsymbol{r}(t) = \boldsymbol{a} \qquad (15.7)$$

в том и только том случае, когда

$$\lim_{t \to t_0} x(t) = a_1, \quad \lim_{t \to t_0} y(t) = a_2, \quad \lim_{t \to t_0} z(t) = a_3. \qquad (15.8)$$

Действительно, в силу соотношений (15.5) и (15.6), для того чтобы выполнялось условие (15.1), необходимо и достаточно, чтобы

$$\lim_{t \to t_0} |x(t) - a_1| = 0, \quad \lim_{t \to t_0} |y(t) - a_2| = 0, \quad \lim_{t \to t_0} |z(t) - a_3| = 0.$$

Определение 2. *Если t_0 — конечная точка (т. е. число) и для функции $\boldsymbol{r}(t)$, $t \in X$, имеет место равенство*

$$\lim_{t \to t_0} \boldsymbol{r}(t) = \boldsymbol{r}(t_0), \qquad (15.9)$$

то эта функция называется непрерывной в точке t_0.

Как и в случае скалярных функций, условие (15.9) выполняется тогда и только тогда, когда существует предел $\lim\limits_{t \to t_0} \boldsymbol{r}(t)$ и точка t_0 принадлежит множеству X. Из эквивалентности условий (15.7) и (15.8) следует, что векторная функция непрерывна в некоторой точке тогда и только тогда, когда в этой точке непрерывны все ее координатные функции.

Отметим некоторые **свойства пределов векторных функций.**

1^0. *Если* $\lim\limits_{t \to t_0} \boldsymbol{r}(t) = \boldsymbol{a}$, *то* $\lim\limits_{t \to t_0} |\boldsymbol{r}(t)| = |\boldsymbol{a}|$.

Это непосредственно следует из неравенства $\big||\boldsymbol{r}| - |\boldsymbol{a}|\big| \leqslant |\boldsymbol{r} - \boldsymbol{a}|$.

Геометрический смысл этого неравенства состоит в том, что разность длин двух сторон треугольника не превышает длины его третьей стороны.

2^0. $\lim\limits_{t \to t_0} [\boldsymbol{r}_1(t) + \boldsymbol{r}_2(t)] = \lim\limits_{t \to t_0} \boldsymbol{r}_1(t) + \lim\limits_{t \to t_0} \boldsymbol{r}_2(t)$.

3^0. $\lim\limits_{t \to t_0} f(t)\,\boldsymbol{r}(t) = \lim\limits_{t \to t_0} f(t) \lim\limits_{t \to t_0} \boldsymbol{r}(t)$ ($f(t)$ — *скалярная функция*).

4^0. $\lim\limits_{t \to t_0} \boldsymbol{r}_1(t)\,\boldsymbol{r}_2(t) = \lim\limits_{t \to t_0} \boldsymbol{r}_1(t) \lim\limits_{t \to t_0} \boldsymbol{r}_2(t)$.

5^0. $\lim\limits_{t \to t_0} \boldsymbol{r}_1(t) \times \boldsymbol{r}_2(t) = \lim\limits_{t \to t_0} \boldsymbol{r}_1(t) \times \lim\limits_{t \to t_0} \boldsymbol{r}_2(t)$.

В свойствах 1^0—5^0 все рассматриваемые функции определены на некотором множестве $X \subset \boldsymbol{R}$. В свойствах 2^0—5^0 предполагается, что все пределы, входящие в правые части равенств, существуют и утверждается, что существуют пределы, стоящие в левых частях равенств, причем имеют место написанные формулы.

Все эти свойства доказываются аналогично тому, как были доказаны подобные утверждения, встречавшиеся раньше (см. п. 4.9, 5.10). Докажем, например, свойство 5^0. Предварительно заметим, что для любых векторов $\boldsymbol{p}$ и $\boldsymbol{q}$

$$|\boldsymbol{p} \times \boldsymbol{q}| = |\boldsymbol{p}|\,|\boldsymbol{q}| \sin \widehat{\boldsymbol{p}\boldsymbol{q}} \leqslant |\boldsymbol{p}|\,|\boldsymbol{q}|. \tag{15.10}$$

Поэтому если $\boldsymbol{p} = \boldsymbol{p}(t)$, $\boldsymbol{q} = \boldsymbol{q}(t)$, причем $\lim\limits_{t \to t_0} |\boldsymbol{p}(t)| = 0$, а $|\boldsymbol{q}(t)|$ — ограниченная функция, то из (15.10) имеем (см. п. 5.11)

$$\lim_{t \to t_0} |\boldsymbol{p} \times \boldsymbol{q}| = 0. \tag{15.11}$$

Пусть теперь $\lim\limits_{t \to t_0} \boldsymbol{r}_1(t) = \boldsymbol{a}$, $\lim\limits_{t \to t_0} \boldsymbol{r}_2(t) = \boldsymbol{b}$. Положим $\boldsymbol{\alpha}(t) = \boldsymbol{r}_1(t) - \boldsymbol{a}$, $\boldsymbol{\beta}(t) = \boldsymbol{r}_2(t) - \boldsymbol{b}$; тогда, согласно (15.1),

$$\lim_{t \to t_0} |\boldsymbol{\alpha}(t)| = \lim_{t \to t_0} |\boldsymbol{\beta}(t)| = 0 \tag{15.12}$$

и

$$\boldsymbol{r}_1(t) \times \boldsymbol{r}_2(t) = [\boldsymbol{a} + \boldsymbol{\alpha}(t)] \times [\boldsymbol{b} + \boldsymbol{\beta}(t)] =$$
$$= \boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{a} \times \boldsymbol{\beta}(t) + \boldsymbol{\alpha}(t) \times \boldsymbol{b} + \boldsymbol{\alpha}(t) \times \boldsymbol{\beta}(t),$$

где, в силу (15.11),

$$\lim_{t \to t_0} |\boldsymbol{a} \times \beta(t)| = \lim_{t \to t_0} |\alpha(t) \times \boldsymbol{b}| = \lim_{t \to t_0} |\alpha(t) \times \beta(t)| = 0,$$

а так как

$$|\boldsymbol{a} \times \beta(t) + \alpha(t) \times \boldsymbol{b} + \alpha(t) \times \beta(t)| \leqslant$$
$$\leqslant |\boldsymbol{a} \times \beta(t)| + |\alpha(t) \times \boldsymbol{b}| + |\alpha(t) \times \beta(t)|,$$

то и $\lim\limits_{t \to t_0} |\boldsymbol{a} \times \beta(t) + \alpha(t) \times \boldsymbol{b} + \alpha(t) \times \beta(t)| = 0$. Поэтому

$$\lim_{t \to t_0} |\boldsymbol{r}_1(t) \times \boldsymbol{r}_2(t) - \boldsymbol{a} \times \boldsymbol{b}| = 0,$$

а это означает (см. определение 1), что

$$\lim_{t \to t_0} \boldsymbol{r}_1(t) \times \boldsymbol{r}_2(t) = \boldsymbol{a} \times \boldsymbol{b},$$

т. е. свойство 5^0 доказано. □

Отметим, что свойства 1^0—5^0 пределов векторных функций можно, конечно, получить с помощью формул (15.5) из соответствующих свойств скалярных функций, если перейти к координатной записи векторов и их скалярных и векторных произведений.

Из свойств пределов векторных функций следует, что сумма, скалярное и векторное произведения векторных функций, а также произведение скалярных функций на векторные являются непрерывными в некоторой точке, если в этой точке непрерывны все слагаемые, соответственно сомножители.

15.2. Производная и дифференциал векторной функции

Всюду в дальнейшем в настоящем параграфе будет предполагаться, что t_0 — число, т. е. что t_0 не является бесконечно удаленной точкой прямой.

Пусть векторная функция $\boldsymbol{r}(t)$ определена в некоторой окрестности точки t_0; тогда отношение $\frac{\boldsymbol{r}(t) - \boldsymbol{r}(t_0)}{t - t_0}$ определено в соответствующей проколотой окрестности этой точки.

Определение 3. *Предел* $\lim\limits_{t \to t_0} \frac{\boldsymbol{r}(t) - \boldsymbol{r}(t_0)}{t - t_0}$ (*если он, конечно, существует*) *называется производной данной векторной функции в точке* t_0 *и обозначается* $\boldsymbol{r}'(t_0)$ *или* $\dot{\boldsymbol{r}}(t_0)$.

Таким образом, производная векторной функции в точке есть вектор.

Если положить $\Delta t = t - t_0$, $\Delta \boldsymbol{r} = \boldsymbol{r}(t_0 + \Delta t) - \boldsymbol{r}(t_0)$, то

$$\boldsymbol{r}'(t_0) = \lim_{\Delta t \to t_0} \frac{\Delta \boldsymbol{r}}{\Delta t}. \qquad (15.13)$$

Для того чтобы функция $\boldsymbol{r}(t) = (x(t),\ y(t),\ z(t))$, определенная в некоторой окрестности точки t_0, имела производную в t_0, необходимо и достаточно, чтобы функции $x(t)$, $y(t)$ и $z(t)$ имели производные при $t = t_0$, причем в этом случае

$$\boldsymbol{r}'(t_0) = (x'(t_0),\ y'(t_0), z'(t_0)),$$

$$|\boldsymbol{r}'(t_0)| = \sqrt{(x')^2(t_0) + y'^2(t_0) + z'^2(t_0)}\,.$$

Это непосредственно следует из эквивалентности двух подходов (15.7) и (15.8) к определению предела для векторной функции:

$$\lim_{\Delta t \to 0} \frac{\boldsymbol{r}(t_0 - \Delta t) - \boldsymbol{r}(t_0)}{\Delta t} = \left(\lim_{\Delta t \to 0} \frac{x(t_0 + \Delta t) - x(t_0)}{\Delta t},\right.$$

$$\left. \lim_{\Delta t \to 0} \frac{y(t_0 + \Delta t) - y(t_0)}{\Delta t},\ \lim_{\Delta t \to 0} \frac{z(t_0 + \Delta t) - z(t_0)}{\Delta t} \right).$$

Производную $\boldsymbol{r}'(t)$ векторной функции $\boldsymbol{r}(t)$ называют также *скоростью изменения вектора* $\mathbf{r}(t)$ *относительно параметра* t. В том случае, когда длины векторов $\boldsymbol{r}(t)$ не меняются, производная $\boldsymbol{r}'(t)$ называется также и *скоростью вращения* вектора $\boldsymbol{r}(t)$, а ее абсолютная величина — *числовым значением скорости* его *вращения*.

По аналогии со случаем скалярных функций, векторную функцию $\boldsymbol{\alpha}(t)$, $t \in X$, называют *бесконечно малой относительно скалярной функции* $\beta(t)$, $t \in X$, при $t \to 0$, и пишут $\boldsymbol{\alpha}(t) = \boldsymbol{o}(\beta(t))$, $t \to 0$, если существует такая векторная функция $\boldsymbol{\varepsilon}(t)$, определенная на том же множестве X, что и функции $\boldsymbol{\alpha}(t)$, $\beta(t)$, для которых в векторной окрестности точки $t = t_0$ имеет место равенство $\boldsymbol{\alpha}(t) = \boldsymbol{\varepsilon}(t)\,\beta(t)$, $t \in X$, и $\lim\limits_{\Delta t \to t_0} \boldsymbol{\varepsilon}(t) = 0$.

Как и для скалярных функций, если $t_0 \in X$, то функция $\boldsymbol{\varepsilon}(t)$ непрерывна в точке t_0 и поэтому $\varepsilon(t_0) = 0$.

Векторная функция аргумента t называется *линейной*, если она имеет вид $\boldsymbol{a}t + \boldsymbol{b}$, где $\boldsymbol{a}$ и $\boldsymbol{b}$ — какие-либо два фиксированных вектора.

После вводных замечаний можно определить понятие дифференцируемости и дифференциала векторной функции.

Определение 4. *Векторная функция* $\boldsymbol{r}(t)$, *определенная в некоторой окрестности точки* t_0, *называется дифференцируемой при* $t = t_0$, *если ее приращение*

$$\Delta \boldsymbol{r} = \boldsymbol{r}(t_0 + \Delta t) - \boldsymbol{r}(t_0)$$

в точке t_0 *представимо в виде*

$$\Delta \boldsymbol{r} = \boldsymbol{a}\Delta t + \boldsymbol{o}(\Delta t), \ \ \Delta t \to 0. \tag{15.14}$$

При этом линейная векторная функция $\boldsymbol{a}\Delta t$ приращения аргумента Δt называется *дифференциалом векторной функции* $\boldsymbol{r}(t)$ в точке t_0 и обозначается через $d\boldsymbol{r}$, т. е. $d\boldsymbol{r} = \boldsymbol{a}\Delta t$.

Таким образом,

$$\Delta \boldsymbol{r} = d\boldsymbol{r} + \boldsymbol{o}(\Delta t), \ \ \Delta t \to 0. \tag{15.15}$$

Здесь функция $\boldsymbol{o}(\Delta t)$ определена при $\Delta t = 0$; она также равна нулю:

$$\boldsymbol{o}(\Delta t)\big|_{\Delta t = 0} \underset{(15{,}14)}{=} (\Delta \boldsymbol{r} - \boldsymbol{a}\Delta t)\big|_{\Delta t = 0} = 0.$$

Следовательно, если представить эту функцию $\boldsymbol{o}(\Delta t)$ в виде

$$\boldsymbol{o}(\Delta t) = \boldsymbol{\varepsilon}(\Delta t)\Delta t,$$

то функция $\boldsymbol{\varepsilon}(\Delta t)$ будет также определена при $\Delta t = 0$, поэтому здесь предел

$$\lim_{\Delta t \to 0} \varepsilon(\Delta t) = 0 \tag{15.16}$$

рассматривается не по проколотой, а по целой окрестности точки $\Delta t = 0$; как было отмечено выше, в этом случае $\varepsilon(0) = 0$.

Из (15.14) очевидным образом следует, что если векторная функция $\boldsymbol{r}(t)$ дифференцируема в точке t_0, то она и непрерывна в этой точке.

Далее, из (15.14) имеем

$$\lim_{\Delta t \to 0} \frac{\Delta \boldsymbol{r}}{\Delta t} = \lim_{\Delta t \to 0} \left[\boldsymbol{a} + \frac{\boldsymbol{o}(\Delta t)}{\Delta t}\right] = \boldsymbol{a},$$

т. е. если векторная функция $\boldsymbol{r}(t)$ дифференцируема в точке t_0, то она имеет в этой точке производную и

$$\boldsymbol{r}'(t_0) = \boldsymbol{a}.$$

Наоборот, если существует производная $\boldsymbol{r}'(t_0) = \lim\limits_{\Delta t \to 0} \frac{\Delta \boldsymbol{r}}{\Delta t}$ и, следовательно,

$$\boldsymbol{r}'(t_0) = \frac{\Delta \boldsymbol{r}}{\Delta t} + \boldsymbol{\varepsilon}(\Delta t),\ \Delta t \neq 0,$$

где

$$\lim_{\substack{\Delta t \to 0 \\ \Delta t \neq 0}} \boldsymbol{\varepsilon}(\Delta t) = 0, \qquad (15.17)$$

то

$$\Delta \boldsymbol{r} = \boldsymbol{r}'(t_0)\Delta t + \boldsymbol{\varepsilon}(\Delta t)\Delta t. \qquad (15.18)$$

Если положить $\varepsilon(0) = 0$, то, в силу выполнения условия (15.17), равенство (15.18) равносильно равенству (15.14) при $\boldsymbol{a} = \boldsymbol{r}'(t_0)$. Таким образом, функция $r(t)$ дифференцируема в точке t_0 и $d\boldsymbol{r}(t_0) = \boldsymbol{r}'(t_0)\Delta t$.

Полагая по определению и в этом случае $dt = \Delta t$, имеем (опуская для простоты обозначение аргумента) $d\boldsymbol{r} = \boldsymbol{r}'dt$ или $\boldsymbol{r}' = \frac{d\boldsymbol{r}}{dt}$.

Пусть теперь $t = t(\tau)$ — дифференцируемая в точке $\tau = \tau_0$ функция, $t_0 = t(\tau_0)$ и $\Delta\tau = \tau - \tau_0$. Из соотношения (15.18) следует, что

$$\frac{\Delta \boldsymbol{r}}{\Delta \tau} = \boldsymbol{r}'_t \frac{\Delta t}{\Delta \tau} + \varepsilon(\Delta t)\frac{\Delta t}{\Delta \tau}, \qquad (15.19)$$

где $\boldsymbol{r}'_t = \boldsymbol{r}'(t_0)$.

Функция $t = t(\tau)$ дифференцируема в точке τ_0, следовательно, она непрерывна в ней, т. е. имеет место равенство $\lim\limits_{\Delta\tau \to 0} \Delta t = 0$. Поэтому из (15.16) получаем, что $\lim\limits_{\Delta\tau \to 0} \varepsilon(\Delta t) = 0$. Кроме того, в точке τ_0 существует конечная производная $t'(\tau_0) = \lim\limits_{\Delta\tau \to 0} \frac{\Delta t}{\Delta \tau}$. В силу всего этого, в точке τ_0 существует производная векторной функции $\boldsymbol{r}(t(\tau))$:

$$\boldsymbol{r}'_\tau = \lim_{\Delta\tau \to 0} \frac{\Delta \boldsymbol{r}}{\Delta \tau} \underset{(15.19)}{=} \lim_{\Delta\tau \to 0} \left(r'_t \frac{\Delta r}{\Delta \tau} + \varepsilon(\Delta t)\frac{\Delta r}{\Delta \tau} \right) = r'_t t'_\tau,$$

или, в другой записи,

$$\frac{d\boldsymbol{r}}{d\tau} = \frac{d\boldsymbol{r}}{dt}\frac{dt}{d\tau}.$$

Из этой формулы, аналогично случаю скалярных функций, вытекает инвариантность записи дифференциала векторной функции: как для зависимой переменной t, так и для

независимой τ имеем $d\boldsymbol{r} = \boldsymbol{r}'_t\, dt$, $d\boldsymbol{r} = \boldsymbol{r}'_\tau\, d\tau$; чтобы из второй формулы получить первую, надо подставить во вторую формулу $\boldsymbol{r}'_\tau = \boldsymbol{r}'_t t'_\tau$ и заметить, что $t'_\tau\, d\tau = dt$.

Приведем **формулы дифференцирования векторной функции** (аргумент для простоты обозначений опущен):

1. $(\boldsymbol{r}_1 + \boldsymbol{r}_2)' = \boldsymbol{r}'_1 + \boldsymbol{r}'_2$.

2. $(f\boldsymbol{r})' = f'\boldsymbol{r} + f\boldsymbol{r}'$.

3. $(\boldsymbol{r}_1\boldsymbol{r}_2)' = \boldsymbol{r}'_1\boldsymbol{r}_2 + \boldsymbol{r}_1\boldsymbol{r}'_2$.

4. $(\boldsymbol{r}_1 \times \boldsymbol{r}_2)' = \boldsymbol{r}'_1 \times \boldsymbol{r}_2 + \boldsymbol{r}_1 \times \boldsymbol{r}'_2$.

Здесь все рассматриваемые функции определены в некоторой окрестности точки t_0 и предполагается, что все производные в правой части каждого равенства существуют при $t = t_0$; тогда в точке t_0 существуют и производные, стоящие в левой части равенства, причем справедливы написанные формулы.

Все эти формулы доказываются аналогично случаю дифференцирования скалярных функций (см. п. 9.5). Докажем, например, формулу 4.

Использовав свойства 2^0 и 5^0 пределов векторных функций, получим

$$(\boldsymbol{r}_1(t) \times \boldsymbol{r}_2(t))'_{t = t_0} = \lim_{\Delta t \to 0} \frac{\boldsymbol{r}_1(t_0 + \Delta t) \times \boldsymbol{r}_2(t_0 + \Delta t) - \boldsymbol{r}_1(t_0) \times \boldsymbol{r}_2(t_0)}{\Delta t} =$$

$$= \lim_{\Delta t \to 0} \left[\frac{\boldsymbol{r}_1(t_0 + \Delta t) - \boldsymbol{r}_1(t_0)}{\Delta t} \times \boldsymbol{r}_2(t_0 + \Delta t) +\right.$$

$$\left.+\, \boldsymbol{r}_1(t_0) \times \frac{r_2(t_0 + \Delta t) - r_2(t_0)}{\Delta t}\right] = \boldsymbol{r}'_1(t_0) \times \boldsymbol{r}_2(t_0) + \boldsymbol{r}_1(t_0) \times r'_2(t_0).$$

Производные высших порядков для векторной функции определяются по индукции: если у векторной функции $\boldsymbol{r}(t)$ в некоторой окрестности точки t_0 задана производная $\boldsymbol{r}^{(n)}(t)$ порядка n, $n = 0, 1, 2, \ldots$ (по определению полагается, что $(\boldsymbol{r}^{(0)})(t) \overset{\text{def}}{=} \boldsymbol{r}(t)$), то производная порядка $n + 1$ в этой точке (если эта производная, конечно, существует) определяется по формуле

$$\boldsymbol{r}^{(n+1)}(t_0) = (\boldsymbol{r}^{(n)}(t))'\big|_{t = t_0}.$$

Если векторная функция $\boldsymbol{r}(t) = (x(t), y(t), z(t))$ определена в некоторой окрестности точки t_0 и имеет n производных в этой точке, то для нее справедлива формула Тейлора

$$\Delta \boldsymbol{r} = \boldsymbol{r}(t_0 + \Delta t) - \boldsymbol{r}(t_0) = \sum_{k=1}^{n} \frac{1}{k!} \frac{d^k \boldsymbol{r}(t_0)}{dt} \Delta t^k + \boldsymbol{o}(\Delta t^n).$$

Она непосредственно следует из разложения по формуле Тейлора координатных функций $x(t)$, $y(t)$ и $z(t)$.

Мы видим, что многие факты, установленные в теории скалярных функций, дословно переносятся на векторные функции. Однако было бы ошибкой думать, что это всегда так: например, в определенном смысле аналог формулы конечных приращений не имеет места для векторных функций.

Действительно, рассмотрим векторную функцию $\boldsymbol{r}(t) = (\cos t, \sin t)$, $0 \leqslant t \leqslant 2\pi$. Так как $\boldsymbol{r}'(t) = (-\sin t, \cos t)$, то $|\boldsymbol{r}'(t)| = \sqrt{\sin^2 t + \cos^2 t} = 1$ при любом $t \in [0, 2\pi]$. Следовательно, в этом случае не существует такой точки $\xi \in [0, 2\pi]$, для которой было бы справедливо равенство, аналогичное формуле конечных приращений Лагранжа для скалярных функций,

$$\boldsymbol{r}(2\pi) - \boldsymbol{r}(0) = 2\pi \boldsymbol{r}'(\xi),$$

так как в левой части равенства стоит нулевой вектор (поскольку $\boldsymbol{r}(2\pi) = \boldsymbol{r}(0)$), а справа — ненулевой: $|\boldsymbol{r}'(\xi)| = 1$.

Некоторой заменой формулы конечных приращений для векторных функций является следующее утверждение.

ТЕОРЕМА 1. *Пусть векторная функция $\boldsymbol{r}(t)$ непрерывна на отрезке $[a, b]$ и дифференцируема внутри него. Тогда существует такая точка $\xi \in (a, b)$, что*

$$|\boldsymbol{r}(b) - \boldsymbol{r}(a)| \leqslant (b - a)\,|\boldsymbol{r}'(\xi)|. \qquad (15.20)$$

Доказательство. Если $\boldsymbol{r}(a) = \boldsymbol{r}(b)$, то неравенство (15.20) справедливо при любом выборе точки $\xi \in (a, b)$, так как его левая часть обращается в нуль.

Пусть $\boldsymbol{r}(a) \neq \boldsymbol{r}(b)$. Оценим длину $|\boldsymbol{r}(b) - \boldsymbol{r}(a)|$ вектора $\boldsymbol{r}(b) - \boldsymbol{r}(a) \neq 0$. Если задан какой-либо вектор $\boldsymbol{x}$, то, обозначая через $\boldsymbol{e}$ единичный вектор в направлении вектора $\boldsymbol{x}$, получим $|\boldsymbol{x}| = (\boldsymbol{x}, \boldsymbol{e})$, ибо, согласно определению скалярного произведения,

$$(\boldsymbol{x}, \boldsymbol{e}) = |\boldsymbol{x}|\,|\boldsymbol{e}| \cos \widehat{\boldsymbol{x}\boldsymbol{e}},$$

где $|e| = 1$, $\widehat{\boldsymbol{x}\boldsymbol{e}} = 0$, и, следовательно, $\cos \widehat{\boldsymbol{x}\boldsymbol{e}} = 1$. Поэтому, если e — единичный вектор в направлении вектора $r(b) - r(a) \neq 0$, то

$$|\boldsymbol{r}(b) - \boldsymbol{r}(a)| = (\boldsymbol{r}(b) - \boldsymbol{r}(a), \boldsymbol{e}) = (\boldsymbol{r}(b), \boldsymbol{e}) - (\boldsymbol{r}(a), \boldsymbol{e}),$$

т. е. получилась разность значений числовой функции

$$f(t) = (\boldsymbol{r}(t), \boldsymbol{e}) \tag{15.21}$$

на концах отрезка $[a, b]$:

$$|\boldsymbol{r}(b) - \boldsymbol{r}(a)| = f(b) - f(a). \tag{15.22}$$

Из (15.21) следует, что функция $f(t)$ непрерывна на отрезке $[a, b]$ и дифференцируема во всех его внутренних точках, так как, согласно условиям теоремы, этими свойствами обладает функция $\boldsymbol{r}(t)$. Поэтому, в силу формулы конечных приращений Лагранжа, существует такая точка $\xi \in (a, b)$, что

$$f(b) - f(a) = f'(\xi)(b - a).$$

Но, согласно правилу дифференцирования скалярного произведения, имеем

$$f'(t) = (\boldsymbol{r}'(t), \boldsymbol{e}),$$

откуда

$$f(b) - f(a) = (\boldsymbol{r}'(\xi), \boldsymbol{e})(b - a),\ a < \xi < b. \tag{15.23}$$

Для любых двух векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ из определения скалярного произведения следует неравенство

$$(\boldsymbol{x}, \boldsymbol{y}) \leqslant |(\boldsymbol{x}, \boldsymbol{y})| = |\boldsymbol{x}|\,|\boldsymbol{y}|\,|\cos \widehat{\boldsymbol{x}\boldsymbol{y}}| \leqslant |\boldsymbol{x}|\,|\boldsymbol{y}|;$$

в частности,

$$(\boldsymbol{r}'(\xi), \boldsymbol{e}) \leqslant |\boldsymbol{r}'(\xi)|\,|\boldsymbol{e}| = |\boldsymbol{r}'(\xi)|.$$

Следовательно, из (15.23) получаем

$$f(b) - f(a) \leqslant |\boldsymbol{r}'(\xi)|\,(b - a),\ \ a < \xi < b.$$

Из этого неравенства и формулы (15.22) сразу следует неравенство (15.20). □

§ 16.

Длина кривой

16.1. Понятие пути

Рассмотрим отображения отрезков в трехмерное пространство $\boldsymbol{R}^3$. Пусть $[a, b]$ — некоторый отрезок, а $r(t)$ — его отображение в $\boldsymbol{R}^3$, т. е. отображение, ставящее в соответствие каждой точке $t \in [a, b]$ точку $r(t)$ пространства $\boldsymbol{R}^3$, короче — r: $[a, b] \to \boldsymbol{R}^3$.

Будем считать, что в пространстве $\boldsymbol{R}^3$ фиксирована система координат. В этом случае задание точки пространства равносильно заданию трех ее координат. Обозначим координаты точки $r(t)$ через $x(t)$, $y(t)$, $z(t)$:

$$r(t) = (x(t),\, y(t),\, z(t)). \tag{16.1}$$

Тогда задание отображения $r(t)$ оказывается равносильным заданию трех числовых функций $x(t)$, $y(t)$, $z(t)$, называемых *координатными функциями отображения* $r(t)$.

Понятие предела отображения $r(t)$ определим с помощью его координатных функций. Будем говорить, что отображение $r(t)$ *имеет предел в точке* t_0, *если в этой точке имеют предел все координатные функции, причем*

$$\lim_{t \to t_0} r(t) \overset{\text{def}}{=} \left(\lim_{t \to t_0} x(t),\ \lim_{t \to t_0} y(t),\ \lim_{t \to t_0} z(t)\right). \tag{16.2}$$

Отображение $r(t)$ называется *непрерывным на отрезке* $[a, b]$, *если на этом отрезке непрерывны все его координатные функции*. В этом случае

$$\lim_{t \to t_0} r(t) \underset{(16.2)}{=} \left(\lim_{t \to t_0} x(t),\ \lim_{t \to t_0} y(t),\ \lim_{t \to t_0} z(t)\right) =$$
$$= (x(t_0),\, y(t_0),\, z(t_0)) \underset{(16.2)}{=} r(t_0). \tag{16.3}$$

Для отображения $r(t)$ будем обозначать полужирным шрифтом $\boldsymbol{r}(t)$ векторную функцию, у которой координаты вектора $\boldsymbol{r}(t)$ совпадают с координатами точки $r(t)$, т. е. $\boldsymbol{r}(t) = (x(t),\, y(t),\, z(t))$, и будем называть отображение $r(t)$ и векторную функцию $\boldsymbol{r}(t)$ соответствующими друг другу.

Очевидно, что отображение $r(t)$, $a \leqslant t \leqslant b$, непрерывно на отрезке $[a, b]$ тогда и только тогда, когда на этом отрезке непрерывна соответствующая ему векторная функция $\boldsymbol{r}(t)$. Действительно, векторная функция непрерывна на отрезке в том и только том случае, когда на нем непрерывны все ее координаты (см. п. 15.1), что, по определению, является условием непрерывности отображения $r(t)$ на отрезке.

Определение 1. *Непрерывное отображение отрезка в пространство называется путем, а образ отрезка при рассматриваемом отображении — носителем этого пути.*

Одно и то же множество пространства может быть непрерывным образом отрезков при различных отображениях, т. е. оно может быть носителем разных путей.

Рассматриваемые непрерывные отображения отрезков в пространство не предполагаются взаимно однозначными, поэтому для данного пути

$$r(t) \in \boldsymbol{R}^3, \quad a \leqslant t \leqslant b, \tag{16.4}$$

в одну и ту же точку пространства может отобразиться несколько точек отрезка $[a, b]$. Точки носителя пути (16.4), в которые отображается более одной точки отрезка $[a, b]$, называются *кратными точками* этого пути или *точками* его *самопересечения*.

Для пути (16.4) переменная t называется его *параметром*.

Если точка M носителя пути (16.4) является его кратной точкой, то существует по крайней мере два таких разных значения параметра t_1 и t_2, $a \leqslant t_1 \leqslant b$, $a \leqslant t_2 \leqslant b$, что $r(t_1) = r(t_2) = M$.

Точкой пути (16.4) называется всякая пара $(t, r(t))$, а пространственная точка $r(t) \in \boldsymbol{R}^3$ — носителем этой точки пути. Очевидно, что точка пути (16.4) однозначно определяется значением параметра t.

Точки $(a, r(a))$ и $(b, r(b))$ называются *концами* пути. Точка $(t, r(t))$, для которой значение параметра t лежит в интервале (a, b), называется *внутренней точкой* пути (16.4).

Если $r(a) = r(b)$, то путь называется *замкнутым*, а его носитель — *замкнутым контуром*. Носитель замкнутого пути, не имеющего точек самопересечения, кроме носителей начала и конца пути, называется *простым замкнутым контуром*.

Носитель пути, заданного взаимно однозначным непрерывным отображением отрезка в пространство, называется *простой дугой*.

Будем говорить, что *последовательность* $(t_n, r(t_n))$ *точек* пути (16.4) *стремится по* этому *пути* к его точке $(t_0, r(t_0))$, если

$$\lim_{n \to \infty} t_n = t_0, \quad t_0 \in [a, b], \quad t_n \in [a, b], \quad n = 1, 2, \ldots .$$

В этом случае

$$\lim_{n \to \infty} r(t_n) \underset{(16.3)}{=} r(t_0).$$

Как уже отмечалось, отображение (16.4), т. е. путь, можно задавать в координатном виде, иначе говоря, задать координаты точки $r(t)$ как функции параметра t:

$$r(t) = (x(t), y(t), z(t)), \quad a \leqslant t \leqslant b.$$

В этом случае тройка функций $x(t)$, $y(t)$, $z(t)$, $a \leqslant t \leqslant b$, называется *координатным представлением пути*.

Отображение (16.4) можно задать и соответствующей ему векторной функцией

$$\boldsymbol{r}(t),\ a \leqslant t \leqslant b, \tag{16.5}$$

где, как обычно, вектор $\boldsymbol{r}(t)$ имеет те же координаты, что и точка $r(t)$. Если не оговорено что-либо другое, то всегда предполагается, что $\boldsymbol{r}(t)$ является радиусом-вектором с началом в начале координат. В этом случае носитель пути (16.4) называется годографом векторной функции $\boldsymbol{r}(t)$, а сама векторная функция $\boldsymbol{r}(t)$ — *векторным представлением пути* (16.4).

Если путь лежит в некоторой плоскости, то он называется *плоским*. Если эта плоскость является координатной плоскостью, например плоскостью переменных x и y, то координатное представление пути (16.4) имеет вид

$$x = x(t),\ y = y(t),\ z = 0,\ a \leqslant t \leqslant b,$$

причем уравнение $z = 0$, если это не может привести к недоразумению, обычно не пишут.

Всякая непрерывная на некотором отрезке $[a, b]$ функция $y = f(x)$ является плоским путем, а ее график — носителем этого пути. Рассмотрим конкретные примеры.

Примером замкнутого контура является окружность. Возьмем для определенности окружность радиуса R с центром в начале координат. Ее можно, например, представить как непрерывный образ отрезка $[0, 2\pi]$ с помощью координатных функций

$$x = R\cos t,\ y = R\sin t,\ 0 \leqslant t \leqslant 2\pi. \tag{16.6}$$

Очевидно, окружность является простым замкнутым контуром. Пример незамкнутого пути можно получить, взяв сужение отображения (16.6) на отрезке $[0, \alpha]$, где $0 \leqslant \alpha < 2\pi$.

Отметим, что носитель пути

$$x = R\cos t,\ y = R\sin t,\ 0 \leqslant t < 4\pi, \tag{16.7}$$

совпадает с носителем пути (16.6): и в том и другом случае им является окружность $x^2 + y^2 = R^2$ на плоскости пере-

менных x и y. Однако получена она как результат разных отображений: при отображении (16.6), т. е. при изменении параметра t от 0 до 2π, эта окружность проходится один раз, а при отображении (16.7), т. е. при изменении параметра t от 0 до 4π, она проходится дважды — отображения (16.6) и (16.7) представляют собой разные пути.

Наряду с общим понятием пути нам понадобятся в дальнейшем специальные виды путей: (непрерывно) дифференцируемые, дважды (непрерывно) дифференцируемые и т. п. Определим понятие n раз (непрерывно) дифференцируемого пути. Путь

$$r(t) = (x(t), y(t), z(t)) \in \boldsymbol{R}^3, \ a \leqslant t \leqslant b,$$

называется *n раз (непрерывно) дифференцируемым путем*, если все его координатные функции $x(t)$, $y(t)$, $z(t)$ n раз непрерывно дифференцируемы на отрезке $[a, b]$.

16.2. Понятие кривой

При определении понятия кривой будем исходить из физического представления о траектории пути движущейся в пространстве материальной точки. На такой траектории можно выбирать различные параметры, точно описывающие положение на ней движущейся точки, например время движения t, длину пройденного пути s или что-либо другое. Различным параметрам соответствуют разные отображения отрезков на траекторию, каждое из которых дает ее полное описание.

С геометрической точки зрения иногда два разных пути, например, $x = \cos t$, $y = -\sin t$, $-\pi \leqslant t \leqslant 0$ и $y = \sqrt{1 - x^2}$, $-1 \leqslant x \leqslant 1$, или естественно было бы считать представлением одной и той же «кривой», в данном случае полуокружности $x^2 + y^2 = 1$, $y \geqslant 0$.

В силу всех этих соображений, естественно определить кривую как класс в каком-то смысле равноправных непрерывных отображений отрезков в пространство, т. е. как соответствующий класс путей.

Определение 2. *Путь*

$$r(t),\ a \leqslant t \leqslant b, \qquad (16.8)$$

называется эквивалентным пути

$$\rho(\tau),\ \alpha \leqslant \tau \leqslant \beta, \qquad (16.9)$$

если существует такая непрерывная строго монотонная функция φ (*возрастающая или убывающая*), *отображающая отрезок* $[\alpha, \beta]$ *на отрезок* $[a, b]$, *что для каждого* $\tau \in [\alpha, \beta]$ *справедливо равенство* (рис. 75)

$$r(\varphi(\tau)) = \rho(\tau). \qquad (16.10)$$

Рис. 75

Функция φ называется в этом случае *допустимым преобразованием параметра* τ в параметр t или отображением, осуществляющим эквивалентность пути (16.8) с путем (16.9). Если путь (16.8) эквивалентен пути (16.9), то пишут $r(t) \sim \rho(\tau)$.

Легко убедиться, что каждый путь эквивалентен самому себе: $r(t) \sim r(t)$ (здесь допустимым преобразованием параметра является тождественное отображение $t = \tau$, $a = \alpha \leqslant \tau \leqslant \beta = b$). Это свойство эквивалентности называется ее *рефлексивностью*. Можно проверить также, что если путь (16.8) эквивалентен пути (16.9): $r(t) \sim \rho(t)$, то и путь $\rho(t)$ эквивалентен пути $r(t)$: $\rho(t) \sim r(t)$ (в самом деле, если φ — допустимое преобразование параметра τ в t, то обратная функция φ^{-1} также непрерывна и строго монотонна, следовательно, является допустимым преобразованием параметра, на этот раз t в τ). Это свойство эквивалентности называется ее *симметричностью*.

Наконец, если для трех путей $r_1(t)$, $r_2(t_2)$ и $r_3(t_3)$ имеем $r_1(t_1) \sim r_2(t_2)$ и $r_2(t_2) \sim r_3(t_3)$, то $r_1(t_1) \sim r_3(t_3)$. Это свойство эквивалентности называется ее *транзитивностью* (оно сразу вытекает из того, что композиция непрерывных строго монотонных отображений отрезков также непрерывна и строго монотонна, т. е. композиция допустимых преобразований параметров является допустимым преобразованием).

Если в некотором множестве элементов введено соотношение эквивалентности, обладающее свойствами рефлексивности, симметричности и транзитивности, то такое множество распадается на непересекающиеся классы эквивалентных элементов (см. далее § 63). В данном случае получаются непересекающиеся классы эквивалентных между собой путей.

Отметим еще, что если два пути (16.8) и (16.9) эквивалентны, то их носители совпадают. Это сразу следует из условия (16.10).

Перейдем теперь к определению кривой.

Определение 3. *Всякий класс* Γ *эквивалентных путей называется кривой или, более подробно, непрерывной параметрически заданной кривой.*

Каждый путь из этого класса, т. е. отображение вида (16.4), называется *представлением кривой* Γ, тройка соответствующих координатных функций (16.1) — ее *координатным представлением*, а соответствующая векторная функция — ее *векторным представлением*.

Очевидно, что параметрически заданная кривая однозначно определяется каждым из своих представлений, так как если имеется один путь, то все эквивалентные ему пути получаются с помощью всевозможных допустимых преобразований параметров. Таким образом, для того чтобы задать кривую, надо задать некоторое ее представление.

Кривая Γ, заданная каким-либо своим представлением вида (16.4), (16.1) или (16.5), обозначается соответственно следующим образом:

$$\Gamma = \{r(t);\ a \leqslant t \leqslant b\},$$

$$\Gamma = \{x(t),\ y(t),\ z(t);\ a \leqslant t \leqslant b\}, \qquad (16.11)$$

$$\Gamma = \{\mathbf{r}(t);\ a \leqslant t \leqslant b\}.$$

Примеры. 1. В силу определения эквивалентных путей, два пути, рассмотренные выше, $x = R\cos t$, $y = R\sin t$, $0 \leqslant t \leqslant 2\pi$, и $x = R\cos t$, $y = R\sin t$, $0 \leqslant t \leqslant 4\pi$, не являются эквивалентными (докажите это), но носители их совпадают: они представляют собой одну и ту же окружность $x^2 + y^2 = R^2$.

Два пути $x = \cos t$, $y = \sin t$, $-\pi/2 \leqslant t \leqslant \pi/2$, и $x = \sqrt{\tau(2-\tau)}$, $y = \tau - 1$, $0 \leqslant \tau \leqslant 2$, эквивалентны между собой и поэтому задают одну и ту же кривую. Действительно, функция $\tau = 1 + \sin t$ непрерывна, строго возрастает на отрезке $[-\pi/2, \pi/2]$ и переводит одно представление в другое. Носителями этих путей является полуокружность $x^2 + y^2 = 1$, $x \geqslant 0$.

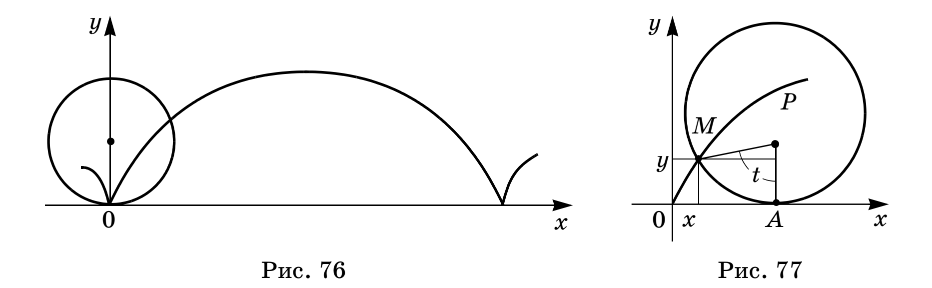


Рис. 76 Рис. 77

2. В качестве второго примера рассмотрим кривую, являющуюся траекторией точки окружности, катящейся без скольжения по прямой. Эта кривая называется *циклоидой*.

Пусть окружность радиуса r катится без скольжения по оси абсцисс в положительном направлении. Пусть в начальный момент времени центр окружности находится в точке $(0; r)$ оси ординат (рис. 76). Напишем уравнение траектории точки $M = (x; y)$ окружности, находящейся в начале качения в начале координат.

В качестве параметра на циклоиде возьмем угол t, на который поворачивается при качении окружность. Пусть, когда окружность повернется на угол t, ее центр окажется в точке P и она будет касаться оси абсцисс в точке A (рис. 77). Тогда $t = \angle MPA$. В силу качения без скольжения длина tr дуги $\overset{\frown}{MA}$ равна длине отрезка OA. Поэтому $x = tr - r \sin t$. Очевидно, $y = r - r \cos t$.

Таким образом, уравнение одной арки циклоиды (т. е. ее части от одного касания окружности с осью абсцисс в фиксированной точке окружности до ближайшего следующего) имеет вид

$$x = r\,(t - \sin t),\ \ y = r\,(1 - \cos t),\ \ 0 \leqslant t \leqslant 2\pi.$$

Мы уже встречались с этими уравнениями, когда рассматривали дифференцируемость функций, заданных параметрически (см. п. 10.3).

Как уже отмечалось выше, два эквивалентных пути имеют один и тот же носитель, поэтому носители всех путей, составляющих кривую, т. е. всех путей, эквивалентных между собой, совпадают.

Определение 4. *Общий носитель всех представлений данной кривой называется носителем этой кривой.*

В случае, когда все пути, составляющие кривую, являются взаимно однозначными отображениями соответствующих отрезков в пространство (для этого достаточно, чтобы этим свойством обладал хотя бы один путь, т. е. хотя бы одно представление кривой) и, следовательно, их носители являются простыми дугами (см. п. 15.1), то и сама кривая называется *простой дугой*.

Определим теперь, что называется точкой параметрически заданной кривой, т. е. кривой, определенной как класс эквивалентных путей.

Определение 5. *Пусть* (16.8) *и* (16.9) — *два представления кривой* Γ, *а* φ — *отображение, осуществляющее их эквивалентность* (см. определение 2): $t = \varphi(\tau)$, $\alpha \leqslant \tau \leqslant \beta$, $a \leqslant t \leqslant b$.

Точки $(t, r(t))$ *и* $(\tau_2, \rho(\tau))$ *соответственно путей* (16.8) *и* (16.9) *называются эквивалентными, если*

$$t = \varphi(\tau). \tag{16.12}$$

Очевидно, что две эквивалентные между собой точки эквивалентных путей имеют один и тот же носитель — это сразу следует из выполнения условий (16.10) и (16.12).

Эквивалентность точек путей будем снова обозначать символом $\sim$:

$$(t, r(t)) \sim (\tau, \varphi(\tau)) \Leftrightarrow t = \varphi(\tau).$$

Легко проверить, что это соотношение эквивалентности также обладает свойством рефлексивности: $(t, r(t)) \sim (t, r(t))$, симметричности: если $(t, r(t) \sim (\tau, \rho(\tau))$, то и $(\tau, \rho(\tau)) \sim (t, r(t))$, транзитивности: если $(t_1, r_1(t_1)) \sim (t_2, r_2(t_2))$ и $(t_2, r_2(t_2)) \sim$ $\sim (t_3, r_3(t_3))$, то $(t_1, r_1(t_1)) \sim (t_3, r_3(t_3))$. Поэтому множество всех точек путей, составляющих заданную кривую, распадается на непересекающиеся классы эквивалентных между собой точек этих путей.

Определение 6. *Каждый класс* $\{(t, r(t))\}$ *эквивалентных между собой точек путей, составляющих некоторую кривую, называется точкой этой кривой, а их общий носитель — носителем этой точки кривой.*

Точка кривой $\{(t, r(t))\}$ называется ее *конечной* или *внутренней точкой*, если эта точка кривой содержит соответственно конечную или внутреннюю точку некоторого (сле-

довательно, и любого) пути, являющегося представлением данной кривой.

Будем говорить, что последовательность точек $\{(t, r(t))\}_n$, $n = 1, 2, \ldots$, кривой Г стремится по ней к ее точке $\{(t, r(t))\}_0$ при $n \to \infty$, если при некотором (а следовательно, и при любом) представлении $r_0(t)$, $a \leqslant t \leqslant b$ кривой Г существуют такие $t_n \in [a, b]$, что $(t_n, r_0(t_n)) \in \{(t, r(t))\}_n$, $n = 1, 2, \ldots$, $(t_0, r_0(t_0)) \in \{(t, r(t))\}_0$ и $\lim\limits_{n \to \infty} t_n = t_0$.

Каждая точка $\{(t, r(t)\}$ кривой Г однозначно определяется каждой отдельной входящей в нее точкой $(t, r(t))$, т. е. соответствующей точкой пути, являющегося одним из представлений кривой Г, а каждая точка $(t, r(t))$ пути (16.8) однозначно определяется значением параметра t. Таким образом, каждая точка кривой Г при выборе какого-либо ее представления $r(t)$, $a \leqslant t \leqslant b$, однозначно определяется значением параметра t. Поэтому точка кривой Г вместо $\{(t, r(t))\}$ обычно просто обозначается через $r(t)$. В силу сказанного, это обозначение имеет однозначный смысл.

Точка носителя кривой называется *кратной* или точкой самопересечения кривой, если она кратная для некоторого, следовательно, и для всякого ее представления.

Очевидно, что совокупность всех носителей точек кривой составляет ее носитель (см. определение 4).

Если кривая является простой дугой, то, так как кратные точки отсутствуют, каждая точка носителя простой дуги однозначно определяет точку кривой, носителем которой она является, и поэтому в данном случае понятие носителя кривой в указанном смысле равносильно понятию кривой.

Как уже было показано на примерах в п. 16.1 (см. (16.6) и (16.7)), неэквивалентные пути могут иметь один и тот же носитель, следовательно, разные кривые также могут иметь один и тот же носитель. Заметим еще, что если $r(a) = r(b)$ при одном представлении кривой, то это условие выполняется и при любом другом ее представлении, т. е. если носитель одного представления кривой является замкнутым контуром, то и носитель всякого другого ее представления также является замкнутым контуром. В этом случае носитель соответствующей кривой также называется *замкнутым контуром*.

Аналогично, только одновременно все носители представлений кривой могут оказаться простыми замкнутыми

контурами, и в этом случае носитель этой кривой также называется *простым замкнутым контуром*.

Перейдем теперь к определению кривых других классов: понятие эквивалентности можно вводить в более узких классах путей и более сильным способом. Это дает возможность определить специальные классы параметрически заданных кривых: n раз дифференцируемых и n раз непрерывно дифференцируемых кривых, $n = 1, 2, \dots$.

Определим для этого сначала соответствующее отношение эквивалентности путей.

Определение 7. *Два* n *раз* (*непрерывно*) *дифференцируемых пути* (см. п. 15.1) *называются* n *раз* (*непрерывно*) *дифференцируемо эквивалентными, если существует функция* φ, *осуществляющая их эквивалентность в смысле определения* 2, *которая как сама, так и ей обратная* n *раз* (*непрерывно*) *дифференцируемы*.

Это отношение эквивалентности также обладает свойствами рефлексивности, симметричности и транзитивности.

Определение 8. *Всякое множество* n *раз* (*непрерывно*) *дифференцируемых и* n *раз* (*непрерывно*) *дифференцируемо эквивалентных между собой путей называется* n *раз* (*непрерывно*) *дифференцируемой параметрически заданной кривой*.

Каждый из путей, входящих в данную кривую, называется ее *представлением* и полностью определяет кривую, т. е. множество всех эквивалентных ему путей в указанном смысле. Функция, осуществляющая n раз (непрерывно) дифференцируемую эквивалентность двух представлений одной и той же n раз (непрерывно) дифференцируемой кривой, называется *допустимым преобразованием параметра*.

Каждая n раз (непрерывно) дифференцируемая кривая содержится как множество путей в некоторой непрерывной кривой, а именно в совокупности всех путей, эквивалентных в смысле определения 2 (тем самым без предположения соответствующей дифференцируемости допустимых преобразований параметров) произвольно выбранному представлению данной кривой. Носитель этой непрерывной кривой называется и носителем исходной n раз (непрерывно) дифференцируемой кривой.

Очевидно, что для n раз (непрерывно) дифференцируемых кривых, как и для непрерывных кривых, их обозначение любым из способов (16.11) имеет однозначный смысл.

16.3. Ориентация кривой. Дуга кривой. Сумма кривых. Неявное задание кривых

Порядок чисел (по величине) на отрезке $[a, b]$ с помощью данного фиксированного представления $r(t)$ кривой $\Gamma\{r(t);\ a \leqslant t \leqslant b\}$, естественно, порождает соответствующий порядок точек на кривой. Точка $r(t') \in \Gamma$ считается предшествующей точке $r(t'') \in \Gamma$, или, что то же, точка $r(t'')$ считается следующей за точкой $r(t')$, если $a \leqslant t' \leqslant t'' \leqslant b$. Точка $r(a)$ называется в этом случае *началом пути*, а точка $r(b)$ — его *концом*. Если указанный порядок точек желательно сохранить и при других представлениях кривой, то необходимо сузить класс допустимых преобразований параметра, а именно, допускать лишь строго возрастающие преобразования параметра.

Два пути (16.8) и (16.9) называются *одинаково ориентированными*, если один из них получается из другого строго возрастающим преобразованием параметра, и противоположно ориентированными, если — строго убывающим преобразованием параметра.

Если дан путь

$$r(t),\ \ a \leqslant t \leqslant b, \tag{16.13}$$

то путь, задаваемый представлением

$$\rho(t) = r(a + b - t),\ \ a \leqslant t \leqslant b, \tag{16.14}$$

является путем, ориентированным противоположно данному.

Все пути, составляющие некоторую кривую, распадаются на два класса таких, что пути каждого из них ориентированы одинаково, а разных — противоположно.

Определение 9. *Для данной кривой каждый из указанных классов называется ориентированной кривой. Каждый из этих классов называется кривой, ориентированной противоположно кривой, определяемой другим классом.*

Если Γ — кривая, то полученные из нее ориентированные кривые обозначаются Γ^+ и Γ^-. Таким образом, если, например, $r(t) \in \Gamma^+$, то

$$r(a + b - t) \in \Gamma^-,\ a \leqslant t \leqslant b. \qquad (16.15)$$

Вместо выражения «задана ориентированная кривая» говорят иногда, что «на кривой задана ориентация» (или порядок точек).

Если кривая имеет концы A и B, то ориентированная кривая, при которой точка A является началом составляющих ее путей, обозначается $\overset{\frown}{AB}$, а при которой B — начало путей, обозначается $\overset{\frown}{BA}$.

Понятие ориентированной кривой имеет смысл, в частности, и для n раз (непрерывно) дифференцируемых кривых.

Если путь $r = r(t)$, $a \leqslant t \leqslant b$, является представлением ориентированной кривой Γ, а функция $\varphi(\tau)$, $\alpha \leqslant \tau \leqslant \beta$, непрерывна, строго убывает и $\varphi(\alpha) = b$, $\varphi(\beta) = a$, то путь $r(\varphi(\tau))$, $\alpha \leqslant \tau \leqslant \beta$, является представлением кривой Γ^-, т. е. противоположно ориентированной кривой.

Если $\tau_0 \in [\alpha, \beta]$ и $t_0 = \varphi(\tau_0)$, то точки $r(t_0)$ и $r(\varphi(\tau_0))$ соответственно кривой Γ и противоположно ориентированной кривой Γ^- называются соответствующими друг другу. Одна точка кривой Γ предшествует другой точке этой кривой тогда и только тогда, когда точка кривой Γ^-, соответствующая первой точке, следует за точкой, соответствующей второй точке. Этим оправдывается термин «противоположно ориентированная кривая».

В заключение сформулируем еще несколько полезных для дальнейшего определений.

Пусть задана кривая $\Gamma = \{r(t);\ a \leqslant t \leqslant b\}$.

Определение 10. *Если* $[a', b'] \subset [a, b]$, *то кривая* $\Gamma' = \{r(t);\ a' \leqslant t \leqslant b'\}$ *называется частью кривой* Γ (*или ее дугой*) *и пишется* $\Gamma' \subset \Gamma$.

Если кривая Γ ориентирована, то ее ориентация порождает ориентацию и на всякой ее части Γ': одна точка кривой Γ' считается следующей за другой ее точкой, если то же самое имеет место для этих точек и на кривой Γ.

Определение 11. *Если* $t_0 \in (a, b)$, $\Gamma_1 = \{r(t);\ a \leqslant t \leqslant t_0\}$, $\Gamma_2 = \{r(t);\ t_0 \leqslant t \leqslant b\}$, *то кривую* $\Gamma = \{r(t);\ a \leqslant t \leqslant b\}$ *называют суммой кривых* Γ_1 *и* Γ_2 *и пишут* $\Gamma = \Gamma_1 \cup \Gamma_2$.

Аналогично определяется сумма конечного числа кривых.

Определение 12. *Сумма конечного числа непрерывно дифференцируемых кривых называется кусочно-непрерывно дифференцируемой кривой.*

Ясно, что кусочно-непрерывно дифференцируемая кривая может просто оказаться непрерывно дифференцируемой кривой.

Определение 13. *Пусть* $\Gamma = \{r(t);\ a \leqslant t \leqslant b\}$ — *плоская кривая, расположенная на плоскости* x, y. *Если существует такая функция* $f(x, y)$, *что координаты* x, y *точек* $r(t)$ *кривой* Γ *удовлетворяют условию*

$$f(x, y) = 0, \tag{16.16}$$

то говорят, что уравнение (16.16) *является неявным представлением кривой* Γ.

Следует, однако, иметь в виду, что, вообще говоря, множество всех точек, удовлетворяющих уравнению вида (16.16), не является кривой в определенном выше смысле даже для достаточно «хороших» функций $f(x, y)$. Например, множество точек, координаты которых удовлетворяют уравнению $(x^2 + y^2)(x^2 + y^2 - 1) = 0$, представляет собой окружность $x^2 + y^2 = 1$ и точку $(0, 0)$. Можно показать, что это множество не является непрерывным образом отрезка.

Можно и в пространственном случае задавать кривые неявным образом, но уже с помощью системы двух уравнений:

$$f_1(x, y, z) = 0,\ \ f_2(x, y, z) = 0.$$

Более подробно этот вопрос будет рассмотрен в п. 41.3.

Наконец, отметим, что кривая всегда ограничена, т. е. лежит в некотором шаре; это следует из того, что функции координатного представления кривой, согласно теореме Больцано—Вейерштрасса, ограничены в силу своей непрерывности. Вместе с тем уже в элементарной математике встречаются неограниченные кривые: прямая, парабола, гипербола, синусоида, график $\operatorname{tg} x$ и т. п. Чтобы охватить и такие «кривые», можно определить класс так называемых *открытых кривых* по схеме, подобной приведенной выше, в которой за основу взято непрерывное отображение интервала, а не отрезка, как это было сделано выше. Открытые кривые, в частности, могут

быть и неограниченными. Подробно и точно сформулировать все эти понятия представляем читателю по мере потребности.

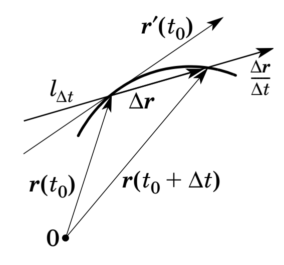


Рис. 78

Иногда кривыми называют и объединение конечного множества кривых в указанном выше смысле. Так, говорят, что гипербола, заданная, например, уравнением $x^2 - y^2 = 1$, является кривой, хотя она состоит из двух непересекающихся «ветвей», каждая из которых представляет собой открытую кривую.

16.4. Касательная к кривой. Геометрический смысл производной векторной функции

Пусть задана кривая $\Gamma = \{\boldsymbol{r}(t);\ a \leqslant t \leqslant b\}$, векторная функция $\boldsymbol{r}(t)$ дифференцируема в точке $t_0 \in [a, b]$ и $\boldsymbol{r}'(t_0) \neq \boldsymbol{0}$. В силу определения дифференцируемости,

$$\Delta \boldsymbol{r} = \boldsymbol{r}(t_0 + \Delta t) - \boldsymbol{r}(t_0) = \boldsymbol{r}'(t_0)\Delta t + \boldsymbol{o}(\Delta t),\ \Delta t \to 0,$$

поэтому из условия $\boldsymbol{r}'(t_0) \neq \boldsymbol{0}$ для всех достаточно малых $\Delta t \neq 0$ имеет место неравенство

$$\boldsymbol{r}(t_0 + \Delta t) \neq \boldsymbol{r}(t_0).$$

Действительно, при сделанных предположениях $\boldsymbol{r}'(t_0)\Delta t \neq \boldsymbol{0}$, поэтому для всех достаточно малых $\Delta t \neq 0$ имеем и

$$\boldsymbol{r}'(t_0)\Delta t + \boldsymbol{o}(\Delta t) \neq \boldsymbol{0}.$$

Прямая, проведенная через точки $r(t_0)$ и $r(t_0 + \Delta t)$, называется *секущей* для кривой Γ. Обозначим ее через $l_{\Delta t}$ (рис. 78). Для всех достаточно малых $\Delta t \neq 0$, в силу условия $r(t_0) \neq r(t_0 + \Delta t)$, секущая $l_{\Delta t}$ определена однозначно. Вектор $\Delta \boldsymbol{r} = \boldsymbol{r}(t_0 + \Delta t) - \boldsymbol{r}(t_0)$ параллелен этой секущей, поэтому вектор $\frac{\Delta r}{\Delta t}$, $\Delta t \neq 0$, отличающийся от вектора $\Delta \boldsymbol{r}$ лишь скалярным множителем $\frac{1}{\Delta t}$, также ей параллелен.

По условию, в точке t_0 существует производная, т. е. предел

$$\lim_{\Delta t \to 0} \frac{\Delta r}{\Delta t} = \boldsymbol{r}'(t_0). \qquad (16.17)$$

Так как все секущие проходят через одну и ту же точку $r(t_0)$, то геометрически формула (16.17) означает, что секущие $l_{\Delta t}$ при $\Delta t \to 0$ стремятся к некоторому предельному положению, т. е. к прямой, проходящей через ту же точку $r(t_0)$ в направлении вектора $\boldsymbol{r}'(t_0)$. Эта прямая, в силу условия $\boldsymbol{r}'(t_0) \neq 0$, определена однозначно. Она и называется *касательной к кривой* Γ в точке $\boldsymbol{r}(t_0)$.

Таким образом, в силу самого определения касательной к кривой Γ в точке $r(t_0)$, производная $\boldsymbol{r}'(t_0)$ векторной функции $\boldsymbol{r}(t)$ в том случае, если $\boldsymbol{r}'(t_0) \neq 0$, является вектором, параллельным касательной в точке $r(t_0)$. Если начало вектора $\boldsymbol{r}'(t_0)$ поместить в эту точку, как это обычно и делается, то он будет направлен по касательной.

В рассматриваемом случае дифференциал $d\boldsymbol{r}(t_0) = \boldsymbol{r}'(t_0)\,dt$ также направлен по касательной к кривой, так как он отличается от производной лишь скалярным множителем dt. Вектор $\boldsymbol{t} = \dfrac{\boldsymbol{r}'}{|\boldsymbol{r}'|}$, $\boldsymbol{r}' \neq 0$, является единичным вектором, направленным по касательной. Вектор $\Delta\boldsymbol{r}$ при $\Delta t > 0$ направлен от точки кривой с меньшим значением параметра к точке с бо́льшим значением параметра, поэтому можно сказать, что вектор $\Delta\boldsymbol{r}$ при $\Delta t > 0$ показывает направление, в котором параметр на кривой возрастает, т. е., как говорят, положительное направление на кривой. Вектор $\dfrac{\Delta\boldsymbol{r}}{\Delta t}$ при $\Delta t > 0$ имеет то же направление, что и вектор $\Delta\boldsymbol{r}$. Поскольку $\lim\limits_{\Delta t \to 0} \dfrac{\Delta\boldsymbol{r}}{\Delta t} = \boldsymbol{r}'(t)$, естественно говорить, что вектор $\boldsymbol{r}'(t)$, следовательно, и вектор $\boldsymbol{t}$, который отличается, быть может, от вектора $\boldsymbol{r}'(t)$ положительным числовым множителем $\dfrac{1}{|\boldsymbol{r}'(t)|}$, также направлены в сторону возрастания параметра и что их ориентация (направление) соответствует ориентации кривой. Направление вектора $\boldsymbol{t}$ (или, что то же, вектора $\boldsymbol{r}'$) называется *положительным направлением касательной*, задаваемым данным ее представлением $\boldsymbol{r}(t)$, $a \leqslant t \leqslant b$.

Уравнение касательной к кривой Γ в точке $r(t_0)$, для которой $\boldsymbol{r}'(t) \neq 0$, в векторной записи имеет вид

$$\boldsymbol{r} = \boldsymbol{r}(t_0) + \boldsymbol{r}'(t_0)\tau, \quad -\infty < \tau < +\infty,$$

где $\boldsymbol{r}$ — текущий радиус-вектор касательной. В координатной записи уравнение касательной в этом случае имеет вид

$$x = x(t_0) + x'(t_0)\tau,\ \ y = y(t_0) + y'(t_0)\tau,\ \ z = z(t_0) + z'(t_0)\tau,$$

$$-\infty < \tau < +\infty.$$

Исключив переменную τ, получим

$$\frac{x - x_0}{x'(t_0)} = \frac{y - y_0}{y'(t_0)} = \frac{z - z_0}{z'(t_0)}.$$

При определении касательной в данной точке дифференцируемой кривой было бы правильнее говорить о касательной данного представления (пути) кривой. Однако если $\boldsymbol{r}(t)$, $a \leqslant t \leqslant b$ является представлением кривой, то всякое другое ее представление имеет вид $\boldsymbol{r}(t(\tau))$, $\alpha \leqslant \tau \leqslant \beta$, где $t = t(\tau)$ — допустимое преобразование параметра, и, следовательно, как сама функция $t(\tau)$, так и ей обратная, являются строго монотонными дифференцируемыми функциями. Поэтому, в силу теоремы 3 п. 9.6 о производной обратной функции, имеем $t'_\tau \tau'_t = 1$, а поэтому для всех $\tau \in [\alpha, \beta]$ выполняется неравенство $t'(\tau) \neq 0$.

Из равенства $\boldsymbol{r}' = \boldsymbol{r}'_\tau t'_\tau$ и условия $t'_\tau \neq 0$ следует, что если один из векторов $\boldsymbol{r}'_t$ и $\boldsymbol{r}'_\tau$ не равен нулю, то не равен нулю и другой и они коллинеарны. Это означает, что вектор, касательный при одном представлении кривой, будет касательным и при другом ее представлении, а поэтому его естественно называть, как это и было сделано, касательным к кривой.

Отметим еще, что если рассматривается ориентированная кривая, то так как в этом случае допустимые преобразования параметра $t(\tau)$, $\alpha \leqslant \tau \leqslant \beta$, строго возрастают, то $t'(\tau) > 0$ во всех точках отрезка $[\alpha, \beta]$. Поэтому из формулы $\boldsymbol{r}' = \boldsymbol{r}'_t t'_\tau$ явствует, что положительное направление касательной одинаково для всех представлений ориентированной кривой, т. е. положительное направление касательной является свойством ориентированной кривой в целом, а не только ее отдельных представлений.

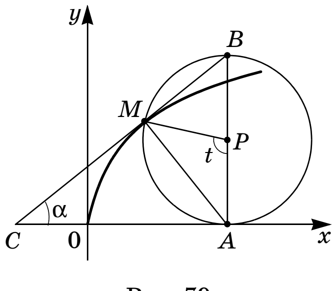


Рис. 79

При преобразовании параметра, меняющего ориентацию кривой, касательный вектор меняет направление на противоположное, так как в этом случае $t'_\tau < 0$.

Пример. Найдем касательные к циклоиде (см. пример 2 в п. 16.2)

$$x = r(t - \sin t),\ \ y = r(1 - \cos t),$$
$$0 \leqslant t \leqslant 2\pi.$$

Вычислим производные: $x' = r(1 - \cos t)$, $y' = r\sin t$. Обозначив через α, $-\frac{\pi}{2} \leqslant \alpha \leqslant \frac{\pi}{2}$, угол, образованный касательной в точке M циклоиды с осью абсцисс, получим

$$\operatorname{tg} \alpha = y'_x = \frac{y'_t}{x'_t} = \frac{\sin t}{1 - \cos t} = \operatorname{ctg} \frac{t}{2} = \operatorname{tg}\left(\frac{\pi}{2} - \frac{t}{2}\right)$$

и, следовательно, $\alpha = \pi/2 - t/2$. Отсюда следует простой способ построения касательных к циклоиде.

Обозначим через B верхнюю точку катящейся окружности, повернувшейся на угол t (рис. 79), тогда $\angle MBA = \frac{1}{2}\,\overset{\frown}{MA} = \frac{t}{2}$. Поэтому, если C — точка пересечения прямой BM с осью абсцисс, то $\angle ACB = \pi/2 - t/2 = \alpha$. Это означает, что прямая CB является касательной к циклоиде.

Итак, касательной к циклоиде в точке M является прямая, соединяющая точку M с верхней точкой B катящейся окружности.

Определение 14. *Пусть* Γ — *дифференцируемая кривая и* $\boldsymbol{r}(t)$, $a \leqslant t \leqslant b$ — *ее векторное представление. Точка* $r(t)$ *кривой* Γ, *в которой* $\boldsymbol{r}'(t) \neq 0$, *называется неособой, а точка, в которой* $\boldsymbol{r}'(t) = 0$, — *особой.*

Выше было показано, что в данной точке кривой при всех представлениях $\boldsymbol{r}(t)$ этой кривой либо одновременно $\boldsymbol{r}' \neq 0$, либо $\boldsymbol{r}' = 0$, поэтому неособая точка при одном представлении дифференцируемой кривой будет неособой и при другом ее представлении. Таким образом, понятие неособой и особой точки не зависит от выбора представления кривой.

Если $\boldsymbol{r}(x(t), y(t), z(t))$, то из равенства $|\boldsymbol{r}'| = \sqrt{x'^2 + y'^2 + z'^2}$ (см. п. 15.2) имеем: *точка* $(x(t), y(t), z(t))$ *кривой* Γ *неособая*

тогда и только тогда, когда в ней ${x'}^2 + {y'}^2 + {z'}^2 > 0$, *т. е. хотя бы одна из производных* x', y' *и* z' *не обращается в нуль.*

Согласно доказанному выше, *во всякой неособой точке кривой* Γ *существует касательная.*

Определение 15. *Непрерывно дифференцируемая кривая без особых точек называется гладкой.*

Кривая, представимая как сумма конечного числа гладких кривых, называется кусочно-гладкой.

Отметим, что если плоская кривая имеет явное представление $y = y(x)$ или $x = x(t)$, то для нее вектор $\boldsymbol{r}'(t) = (x'(t), y'(t))$ всегда не нулевой: в первом случае это $(1, y')$, а во втором — $(x', 1)$.

Аналогично определяется касательная как предельное положение секущей и кривой $\Gamma = \{\boldsymbol{r}(t);\ a \leqslant t \leqslant b\}$ в точке $r(t_0)$, $t_0 \in [a, b]$, и в случае, когда $\boldsymbol{r}'(t_0) = 0$, но существует некоторое натуральное $n > 1$, для которого $\boldsymbol{r}^{(n)}(t_0) \neq 0$.

Если все $\boldsymbol{r}^{(k)}(t_0) = 0$, $k = 1, 2, \ldots, n-1$, а $\boldsymbol{r}^{(n)}(t_0) \neq 0$, то, раскладывая $\Delta\boldsymbol{r}$ по формуле Тейлора, получаем

$$\Delta\boldsymbol{r} = \boldsymbol{r}(t_0 + \Delta t) - \boldsymbol{r}(t_0) = \frac{1}{n!}\boldsymbol{r}^{(n)}(t_0)\Delta t^n + \boldsymbol{o}(\Delta t^n),\ \Delta t \to 0.$$

Вектор $\dfrac{\Delta\boldsymbol{r}}{\Delta t^n}$ направлен параллельно секущей $l_{\Delta t}$, проходящей через точки $r(t_0)$ и $r(t_0 + \Delta t)$. Из написанного равенства следует, очевидно, что существует предел

$$\lim_{\Delta t \to 0} \frac{\Delta\boldsymbol{r}}{\Delta t^n} = \frac{1}{n!}\boldsymbol{r}^{(n)}(t_0) \neq 0.$$

Поэтому в этом случае предельное положение секущей $l_{\Delta t}$, т. е. касательная в точке $r(t_0)$, является прямой, проходящей через точку $r(t_0)$ параллельно вектору $\boldsymbol{r}^{(n)}(t_0)$, и, следовательно, уравнение касательной имеет вид

$$\boldsymbol{r} = \boldsymbol{r}^{(n)}(t_0)\tau + \boldsymbol{r}(t_0),\ -\infty < \tau < +\infty.$$

16.5. Длина кривой

Прежде чем определить понятие длины кривой, введем понятие разбиения отрезка, которое будет неоднократно встречаться в дальнейшем.

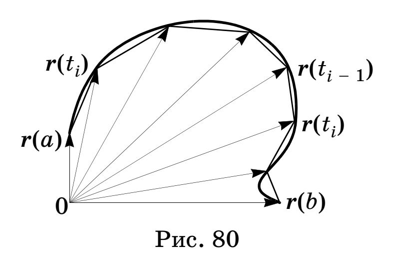


Рис. 80

Определение 16. *Для отрезка* $[a, b]$ *всякую систему его точек* t_i, $i = 0, 1, \ldots, i_\tau$ *таких, что* $a = t_0 < t_1 < \ldots < t_{i-1} < t_i < \ldots < t_{i_\tau} = b$, *будем называть его разбиением и обозначать* $\tau = \{t_i\}_{i=0}^{i=i_\tau}$.

Пусть задана кривая $\Gamma = \{\boldsymbol{r}(t),\ a \leqslant t \leqslant b\}$ и пусть $\tau = \{t_i\}_{i=0}^{i=i_\tau}$ — некоторое разбиение отрезка $[a, b]$. Положим

$$\sigma_\tau \stackrel{\mathrm{def}}{=} \sum_{i=1}^{i_\tau} |\boldsymbol{r}(t_i) - \boldsymbol{r}(t_{i-1})|.$$

Очевидно (рис. 80), σ_τ — длина ломаной с вершинами $r(a)$, $r(t_1)$, ... , $r(t_{i_\tau - 1})$, $r(b)$, т. е., как обычно говорят, *ломаной, вписанной в кривую* Γ.

Всякую ломаную, в частности и вписанную в кривую $\Gamma = \{r(t);\ a \leqslant t \leqslant b\}$, можно рассматривать как кривую в смысле данного выше определения, если только задать ее представление. Пусть Λ — ломаная, т. е. множество, состоящее из конечного числа отрезков с вершинами в точках M_0, M_1, ... , M_n (эти отрезки называются звеньями ломаной). Возьмем некоторый отрезок $[a, b]$ и какое-либо его разбиение на n отрезков: $\tau = \{t_i\}_{i=0}^{i=n}$. Будем для простоты всегда считать, что представлением ломаной является непрерывное отображение $\rho(t)$, линейно отображающее каждый отрезок $[t_{i-1}, t_i]$ на отрезок $M_{i-1}M_i$, $i = 1, 2, \ldots, n$; таким образом, если обозначить через $\boldsymbol{\rho}_i$ радиус-вектор точки M_i, $i = 0, 1, \ldots, n$, то векторное представление ломаной имеет вид

$$\boldsymbol{\rho}(t) = \frac{\rho_i - \rho_{i-1}}{t_i - l_{i-1}}(t - t_{i-1}) + \boldsymbol{\rho}_{i-1},$$

$$t_{i-1} \leqslant t \leqslant t_i,\ i = 1, 2, \ldots, n.$$

Если $M_{i-1} \neq M_i$ при $i = 1, 2, \ldots, n$, то ломаная называется *невырожденной*.

Определение 17. *Для заданной кривой* $\Gamma = \{r(t);\ a \leqslant t \leqslant b\}$ *величина* $S_\Gamma = \sup\limits_\tau \sigma_\tau$, *где верхняя грань взята по всевозмож-*

ным разбиениям τ отрезка $[a, b]$, называется длиной кривой Γ.

Если $S_\Gamma < +\infty$, то кривая Γ называется спрямляемой.

В силу этого определения, спрямляемость кривой и ее длина не зависят от выбора представления кривой и всегда

$$0 \leqslant S_\Gamma \leqslant +\infty.$$

УПРАЖНЕНИЕ 1. Доказать, что кривая, являющаяся частью спрямляемой кривой, также спрямляема.

ЛЕММА 1. *Пусть* $\Gamma = \Gamma_a \cup \Gamma_b$, *тогда*

$$S_\Gamma = S_{\Gamma_a} + S_{\Gamma_b}. \qquad (16.18)$$

Доказательство. Пусть $a < c < b$ и $\Gamma = \{\boldsymbol{r}(t);\ a \leqslant t \leqslant b\}$, $\Gamma_a = \{\boldsymbol{r}(t);\ a \leqslant t \leqslant c\}$, $\Gamma_b = \{\boldsymbol{r}(t);\ c \leqslant t \leqslant b\}$.

Пусть τ — разбиение отрезка $[a, b]$, а τ^* — разбиение этого же отрезка, совпадающее с τ, если точка c входит в разбиение τ, и получающееся из τ добавлением к нему точки c, если эта точка не входит в разбиение τ. Разбиение τ^* является объединением двух разбиений отрезков $[a, c]$ и $[c, b]$, которые мы обозначим соответственно τ_a и τ_b, т. е. $\tau^* = \tau_a \cup \tau_b$. Очевидно, для длин ломаных, соответствующих разбиениям τ^*, τ_a и τ_b, справедливо равенство $\sigma_{\tau^*} = \sigma_{\tau_a} + \sigma_{\tau_b}$. Но $\sup\limits_{\tau_b} \sigma_{\tau_a} = S_{\Gamma a}$, $\sup\limits_{\tau} \sigma_{\tau_b} = S_{\Gamma b}$, следовательно, $\sigma_{\tau^*} \leqslant S_{\Gamma_a} + S_{\Gamma_b}$.

При переходе от разбиения τ к разбиению τ^*, быть может, лишь одно звено $r(t_{i-1})r(t_i)$ заменяется двумя: $r(t_{i-1})r(c)$ и $r(c)r(t_i)$, а так как $|r(t_{i-1})r(t_i)| \leqslant |r(t_{i-1})r(c)| + |r(c)r(t_i)|$, то $\sigma_\tau \leqslant \sigma_{\tau^*}$, следовательно, $\sigma_\tau \leqslant S_{\Gamma_a} + S_{\Gamma_b}$.

Но $S_\Gamma = \sup\limits_{\tau} \sigma_\tau$, поэтому $S_\Gamma \leqslant S_{\Gamma_a} + S_{\Gamma_b}$.

Докажем теперь обратное неравенство. Для произвольных разбиений τ_a и τ_b соответственно отрезков $[a, c]$ и $[c, b]$ и разбиения $\tau^* = \tau_a \cup \tau_b$ отрезка $[a, b]$ имеем $\sigma_{\tau_a} + \sigma_{\tau_b} = \sigma_\tau \leqslant S_\Gamma$. Отсюда $\sigma_{\tau_a} \leqslant S_\Gamma - \sigma_{\tau_b}$; фиксируя разбиение τ_b и переходя к верхней грани σ_{τ_a} при всевозможных τ_a, получаем неравенство $S_{\Gamma_a} \leqslant S_\Gamma - \sigma_{\tau_b}$ и, следовательно, неравенство

$$S_{\Gamma_a} + \sigma_{\tau_b} \leqslant S_\Gamma.$$

Беря верхнюю грань множества чисел σ_{τ_b}, которое получается при всевозможных разбиениях τ_b, имеем $S_{\Gamma_a} + S_{\Gamma_b} \leqslant S_\Gamma$. □

Отметим, что в лемме 1 не предполагается, что рассматриваемые кривые спрямляемы.

Задача 13. Построить пример неспрямляемой кривой.

Докажем одно достаточное условие спрямляемости кривой и получим оценку ее длины.

ТЕОРЕМА 1. *Если кривая* $\Gamma = \{\boldsymbol{r}(t);\ a \leqslant t \leqslant b\}$ *непрерывно дифференцируема, то она спрямляема и ее длина* S_Γ *удовлетворяет неравенствам*

$$|\boldsymbol{r}(b) - \boldsymbol{r}(a)| \leqslant S_\Gamma \leqslant M(b - a), \tag{16.19}$$

где

$$M = \max_{[a,b]} |\boldsymbol{r}'(t)|. \tag{16.20}$$

Отметим, что, в силу непрерывности производной $\boldsymbol{r}'(t)$, ее абсолютная величина $|\boldsymbol{r}'(t)|$ также непрерывна и поэтому достигает на отрезке $[a, b]$ своего наибольшего значения M.

Доказательство. Возьмем какое-либо разбиение

$$\tau = \{t_i\}_{i=0}^{i=i_\tau}$$

отрезка $[a, b]$. Тогда, применив неравенство (15.20), получим

$$|\boldsymbol{r}(b) - \boldsymbol{r}(a)| = \left|\sum_{i=1}^{i_\tau}(\boldsymbol{r}(t_i) - \boldsymbol{r}(t_{i-1}))\right| \leqslant \sum_{i=1}^{i_\tau} |\boldsymbol{r}(t_i) - \boldsymbol{r}(t_{i-1})| \leqslant$$
$$\leqslant \sum_{i=1}^{i_\tau}|\boldsymbol{r}'(\xi_i)|(t_i - t_{i-1}), \tag{16.21}$$

где $\xi_i \in (t_{i-1}, t_i)$, $i = 1, 2, \ldots, i_\tau$.

Так как

$$\sum_{i=1}^{i_\tau}|\boldsymbol{r}(t_i) - \boldsymbol{r}(t_{i-1})| = \sigma_\tau$$

— длина вписанной в кривую Γ ломаной, соответствующей разбиению τ, и для всех $i = 1, 2, \ldots, i_\tau$, в силу (16.20), имеет место неравенство $|\boldsymbol{r}'(\xi_i)| \leqslant M$, то из неравенства (16.21) для любого разбиения τ имеем

$$|\boldsymbol{r}(b) - \boldsymbol{r}(a)| \leqslant \sigma_\tau \leqslant M\sum_{i=1}^{i_\tau}(t_i - t_{i-1}) = M(b - a). \tag{16.22}$$

Перейдя в этом неравенстве к верхней грани по τ, получим утверждение теоремы. □

ТЕОРЕМА 2. *Пусть кривая* $\Gamma = \{\boldsymbol{r}(t) = (x(t), y(t), z(t));\ a \leqslant t \leqslant b\}$ *непрерывно дифференцируема. Тогда переменная длина дуги* s, *отсчитываемая от начала* $\boldsymbol{r}(a)$ *кривой* Γ (*от ее конца* $\boldsymbol{r}(b)$), *является возрастающей* (*убывающей*), *непрерывно дифференцируемой функцией параметра* t; *при этом*

$$\frac{ds}{dt} = \sqrt{x'^2 + y'^2 + z'^2} = \left|\frac{d\boldsymbol{r}}{dt}\right| \tag{16.23}$$

(*соответственно*

$$\left.\frac{ds}{dt} = -\sqrt{x'^2 + y'^2 + z'^2} = -\left|\frac{d\boldsymbol{r}}{dt}\right|\right). \tag{16.24}$$

Доказательство. Пусть $s = s(t)$ — длина дуги кривой Γ от точки $\boldsymbol{r}(a)$ до точки $\boldsymbol{r}(t)$. Пусть $t_0 \in [a, b]$, $t_0 + \Delta t \in [a, b]$ и $\Delta s = s(t_0 + \Delta t) - s(t)$. Очевидно, что функция $s = s(t)$ возрастает на отрезке $[a, b]$, т. е. если $\Delta t > 0$, то $\Delta s \geqslant 0$; если же $\Delta t < 0$, то $\Delta s \leqslant 0$. Поэтому всегда $\frac{\Delta s}{\Delta t} \geqslant 0$.

Применив неравенство (16.19) к части кривой Γ, соответствующей отрезку $[t_0, t_0 + \Delta t]$ при $\Delta t > 0$ или отрезку $[t_0 + \Delta t, t_0]$ при $\Delta t < 0$, получим

$$|\boldsymbol{r}(t_0 + \Delta t) - \boldsymbol{r}(t_0)| \leqslant |\Delta s| \leqslant M|\Delta t|,$$

откуда

$$\left|\frac{\boldsymbol{r}(t_0 + \Delta t) - \boldsymbol{r}(t_0)}{\Delta t}\right| \leqslant \frac{\Delta s}{\Delta t} \leqslant M, \tag{16.25}$$

где M — наибольшее значение $|\boldsymbol{r}'(t)|$ на отрезке $[t_0, t_0 + \Delta t]$ при $\Delta t > 0$ или на отрезке $[t_0 + \Delta t, t_0]$ при $\Delta t < 0$.

В силу непрерывности производной $\boldsymbol{r}'(t)$, ее абсолютная величина $|\boldsymbol{r}'(t)|$ также непрерывна. Поэтому существует наибольшее значение $|\boldsymbol{r}'(t)|$ и оно достигается в некоторой точке $\xi = t_0 + \theta\Delta t$, $0 < \theta < 1$, указанного отрезка. Поэтому неравенство (16.25) можно переписать в виде

$$\left|\frac{\boldsymbol{r}(t_0 + \Delta t) - \boldsymbol{r}(t_0)}{\Delta t}\right| \leqslant \frac{\Delta s}{\Delta t} \leqslant |\boldsymbol{r}'(t_0 + \theta\Delta t|, \quad 0 < \theta < 1.$$

Перейдя здесь к пределу при $\Delta t \to 0$, в левой части неравенства в силу определения производной, а в правой в силу непрерывности производной $\boldsymbol{r}'(t)$ в точке $t = t_0$, получим $|\boldsymbol{r}'(t_0)|$.

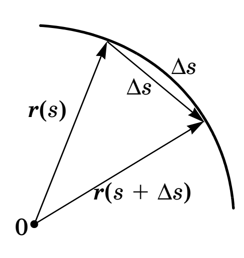


Рис. 81

Следовательно, предел $\lim\limits_{\Delta t \to 0} \dfrac{\Delta s}{\Delta t}$ существует и также равен $|\boldsymbol{r}'(t_0)|$, т. е. существует производная $s'(t_0)$ и имеет место равенство $s'(t_0) = |\boldsymbol{r}'(t_0)|$.

Если $\boldsymbol{r}(t) = (x(t),\ y(t),\ z(t))$, то $\boldsymbol{r}'(t) = (x'(t),\ y'(t),\ z'(t))$ и поэтому

$$s'(t) = |\boldsymbol{r}'(t)| = \sqrt{[x'(t)]^2 + [y'(t)]^2 + [z'(t)]^2}.$$

Если теперь $\sigma = \sigma(t)$ — переменная длина дуги, отсчитываемая от конца $\boldsymbol{r}(b)$ кривой Γ, то, очевидно, $\sigma = S_\Gamma - s$, откуда, дифференцируя это равенство по t, имеем

$$\frac{d\sigma}{dt} = -\frac{ds}{dt} = -\left|\frac{d\boldsymbol{r}}{dt}\right|. \ \square$$

СЛЕДСТВИЕ 1. *Если параметром непрерывно дифференцируемой кривой является переменная длина дуги s, то*

$$\left|\frac{d\boldsymbol{r}}{ds}\right| = 1. \qquad (16.26)$$

Это сразу следует из формулы $\left|\dfrac{d\boldsymbol{r}}{dt}\right| = \dfrac{ds}{dt}$ при $t = s$.

Замечание 1. Формула (16.26) имеет простой геометрический смысл. Поясним его. Пусть параметром непрерывно дифференцируемой кривой Γ является переменная длина дуги s: $\Gamma = \{\boldsymbol{r}(s);\ 0 \leqslant s \leqslant S_\Gamma\}$. Величина $|\Delta\boldsymbol{r}| = |\boldsymbol{r}(s + \Delta s) - \boldsymbol{r}(s)|$ равна длине отрезка, соединяющего точки $\boldsymbol{r}(s)$ и $\boldsymbol{r}(s + \Delta s)$. Этот отрезок называется обычно хордой, стягивающей дугу кривой Γ с началом в точке $\boldsymbol{r}(s)$ и концом в точке $\boldsymbol{r}(s + \Delta s)$. Длина указанной дуги, очевидно, равна $|\Delta s|$ (рис. 81). Так как $\left|\dfrac{d\boldsymbol{r}}{ds}\right| = \lim\limits_{\Delta s \to 0} \left|\dfrac{\Delta\boldsymbol{r}}{\Delta s}\right|$, то из равенства (16.26) следует

$$\lim_{\Delta s \to 0} \frac{|\Delta\boldsymbol{r}|}{|\Delta s|} = 1.$$

Это означает, что предел отношения длины дуги к длине стягивающей ее хорды равен единице, когда дуга стягивается в точку. В этом и состоит геометрический смысл формулы (16.26).

СЛЕДСТВИЕ 2. *Для всякой непрерывно дифференцируемой кривой Γ без особых точек, т. е. для всякой гладкой кривой, существует ее представление $\boldsymbol{r} = \boldsymbol{r}(s)$, в котором за параметр s взята переменная длина дуги кривой Γ.*

Доказательство. Пусть непрерывно дифференцируемая кривая $\Gamma = \{\boldsymbol{r}(t);\ a \leqslant t \leqslant b\}$ не имеет особых точек, т. е. $\boldsymbol{r}'(t) \neq 0$ для всех $t \in [a, b]$. В этом случае переменная длина дуги $s = s(t)$ является строго возрастающей непрерывно дифференцируемой функцией, так как $\frac{ds}{dt} = |\boldsymbol{r}'| > 0$ во всех точках $[a, b]$. Поэтому существует обратная функция $t = t(s)$, $0 \leqslant s \leqslant S_\Gamma$, которая также строго возрастает и имеет непрерывную, не обращающуюся в нуль производную на отрезке $[0, S_\Gamma]$, т. е. функция $t = t(s)$ является допустимым преобразованием параметра для непрерывно дифференцируемых кривых без особых точек и представление $\boldsymbol{r} = \boldsymbol{r}(t(s))$ является искомым представлением, в котором роль параметра играет переменная длина дуги. □

Выясним теперь геометрический смысл координат вектора $\frac{d\boldsymbol{r}}{ds}$. Обозначим через α, β и γ углы, образованные вектором $\frac{d\boldsymbol{r}}{ds}$ или, что то же, касательной к кривой $\Gamma = \{r(s);\ 0 \leqslant s \leqslant S_\Gamma\}$ соответственно с осями Ox, Oy и Oz. Тогда из равенства $\left|\frac{d\boldsymbol{r}}{ds}\right| = 1$, очевидно, следует, что проекции вектора $\frac{d\boldsymbol{r}}{ds}$ на оси координат равны соответственно направляющим косинусам вектора $\frac{d\boldsymbol{r}}{ds}$: $\cos\alpha$, $\cos\beta$ и $\cos\gamma$, т. е.

$$\frac{d\boldsymbol{r}}{ds} = (\cos\alpha, \cos\beta, \cos\gamma). \tag{16.27}$$

Наряду с этим, для векторной функции $\boldsymbol{r}(s) = (x(s), y(s), z(s))$, как для всякой векторной функции (см. п. 15.2), имеем

$$\frac{d\boldsymbol{r}}{ds} = \left(\frac{dx}{ds}, \frac{dy}{ds}, \frac{dz}{ds}\right). \tag{16.28}$$

Сравнивая (16.27) и (16.28), получаем

$$\frac{dx}{ds} = \cos\alpha, \quad \frac{dy}{ds} = \cos\beta, \quad \frac{dz}{ds} = \cos\gamma. \tag{16.29}$$

В качестве примера рассмотрим кривую, называемую *винтовой линией*. Эта кривая задается представлением

$$x = a\cos t, \quad y = a\sin t, \quad z = bt, \quad a^2 + b^2 \neq 0, \quad 0 \leqslant t \leqslant T.$$

Очевидно, что винтовая линия является бесконечно дифференцируемой кривой, и так как

$$x'^2 + y'^2 + z'^2 = a^2\sin^2 t + a^2\cos^2 t + b^2 = a^2 + b^2 \neq 0,$$

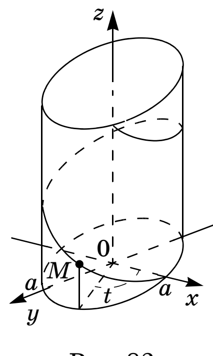


Рис. 82

то она не имеет особых точек (рис. 82). Следовательно, переменную длину ее дуги можно принять за параметр.

Найдем соответствующее представление. Согласно формуле (16.23), имеем

$$\frac{ds}{dt} = \sqrt{x'^2 + y'^2 + z'^2} = \sqrt{a^2 + b^2}.$$

Отсюда $\frac{dt}{ds} = \frac{1}{\sqrt{a^2 + b^2}}$, и так как $t(0) = 0$, то $t = \frac{s}{\sqrt{a^2 + b^2}}$. Поэтому искомое представление имеет вид

$$x(s) = a \cos \frac{s}{\sqrt{a^2 + b^2}}, \quad y(s) = a \sin \frac{s}{\sqrt{a^2 + b^2}},$$

$$z(y) = \frac{bs}{\sqrt{a^2 + b^2}}, \quad 0 \leqslant s \leqslant T\sqrt{a^2 + b^2}.$$

УПРАЖНЕНИЕ 2. Доказать, что для спрямляемой кривой без точек самопересечения переменная длина дуги является непрерывной строго монотонной функцией параметра.

Замечание 2. Если кривая Г гладкая и в качестве параметра на ней выбрана переменная длина дуги s, $0 \leqslant s \leqslant S$, то единичный касательный вектор $\tau = \frac{dr}{ds}$ является непрерывной функцией переменной s. Если какая-то функция является непрерывной функцией параметра кривой, то будем говорить, что она непрерывна вдоль кривой. Теперь можно сказать, что на ориентированной гладкой кривой имеется непрерывный вдоль нее единичный касательный вектор. При изменении ориентации кривой, т. е. при переходе к параметру $s^* = S - s$, касательный вектор меняет свое направление (см. п. 16.4). Действительно, напомним, что так как $\frac{ds}{ds^*} = -1$, то $\tau^* = \frac{d\boldsymbol{r}}{ds^*} = \frac{d\boldsymbol{r}}{ds}\frac{ds}{ds^*} = -\frac{d\boldsymbol{r}}{ds} = -\tau$.

Таким образом, ориентации гладкой кривой однозначным образом соответствует выбор единичного непрерывного касательного вектора $\boldsymbol{\tau}$ или $\boldsymbol{\tau}^*$ вдоль кривой.

Пусть Г — плоская кривая и на плоскости фиксирован базис $\boldsymbol{i}$, $\boldsymbol{j}$. В этом случае каждому касательному вектору τ и

τ^* соответствует единственный перпендикулярный к нему единичный вектор ν, соответственно ν^* (нормаль к кривой) такой, что пары векторов τ, ν и τ^*, ν^* ориентированы так же как пара $\boldsymbol{i}$, $\boldsymbol{j}$ векторов базиса (т. е. так, что определители матриц перехода от векторов τ, ν, соответственно векторов τ^*, ν^*, к базису положительны). Очевидно, что $\nu^* = -\nu$, и нормали ν и ν^* также являются непрерывными функциями вдоль кривой Г. Из сказанного ясно, что ориентация плоской кривой может быть задана не только непрерывным на ней единичным касательным вектором, но и непрерывной вдоль этой кривой единичной нормалью.

16.6. Плоские кривые

Пусть $\Gamma = \{\boldsymbol{r}(t);\ a \leqslant t \leqslant b\}$ — непрерывно дифференцируемая плоская кривая, лежащая в плоскости xOy:

$$\boldsymbol{r}(t) = (x(t),\ y(t)),$$

и пусть $s = s(t)$ — переменная длина дуги кривой Г; для ее производной из формул (16.23) и (16.24) получаем

$$\frac{ds}{dt} = \pm\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}, \tag{16.30}$$

здесь знак плюс берется, если длина дуги $s(t)$ отсчитывается от начальной точки $r(a)$ кривой, и знак минус, если от конечной точки $r(b)$. Из формулы (16.30) для дифференциала дуги получаем выражение

$$ds^2 = dx^2 + dy^2. \tag{16.31}$$

Пусть точка $(x(t_0),\ y(t_0))$ — неособая, т. е. $x'^2(t_0) + y'^2(t_0) > 0$, например $x'(t_0) \neq 0$. Пусть для определенности $x'(t_0) > 0$; тогда в некоторой окрестности точки t_0 также $x'(t) > 0$ и, значит, функция $x(t)$ строго монотонно возрастает в этой окрестности; поэтому существует обратная непрерывно дифференцируемая функция $t = t(x)$. Подставляя ее в представление кривой Г, находим

$$y = y(t(x)) = f(x),$$

т. е. в некоторой окрестности неособой точки непрерывно дифференцируемая кривая является графиком непрерывно дифференцируемой функции f; точнее, существуют окрест-

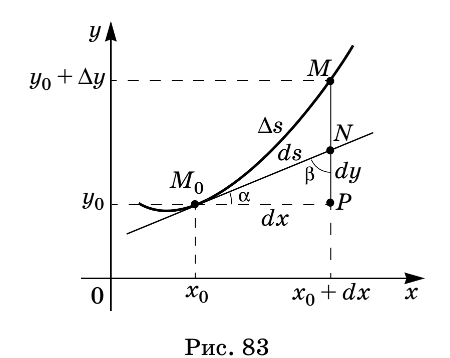


Рис. 83

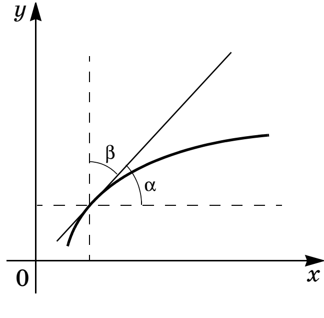


Рис. 84

ность точки t_0 и непрерывно дифференцируемая функция f, определенная на некотором интервале, содержащем точку $x_0 = x(t_0)$, такие, что часть кривой, соответствующая значениям параметра, принадлежащим указанной окрестности точки t_0, является графиком функции f.

В том случае, если кривая Γ является графиком непрерывно дифференцируемой функции $y = f(x)$, $a \leqslant x \leqslant b$, т. е. параметром кривой является переменная x, то $x' = 1$ и поэтому

$$\frac{ds}{dx} = \sqrt{1 + y'^2},$$

следовательно,

$$ds = \sqrt{1 + y'^2}\, dx.$$

Рассмотрим геометрический смысл формулы (16.31) в том случае, когда Γ является графиком непрерывно дифференцируемой функции $y = f(x)$, $a \leqslant x \leqslant b$, и длина дуги кривой отсчитывается от начальной точки кривой (рис. 83). Пусть

$$x_0 \in [a, b], \quad x_0 + dx \in [a, b], \quad y_0 = f(x_0), \quad M_0 = (x_0, y_0),$$

$$y_0 + \Delta y = f(x_0 + dx), \quad M = (x_0 + dx, y_0 + \Delta y),$$

M_0N — касательная в точке M_0; $PM = \Delta y$ — приращение функции в точке $x_0 + dx$; $PN = dy$ — приращение ординаты касательной в точке $x_0 + dx$. Треугольник M_0NP прямоугольный; так как $M_0P = dx$, $PN = dy$, то

$$M_0N^2 = M_0P^2 + PN^2 = dx^2 + dy^2 = ds^2,$$

т. е. длина отрезка касательной M_0N равна ds. Иначе говоря, приращение длины касательной, т. е. $\sqrt{dx^2 + dy^2}$, равно главной части ds приращения длины дуги Δs.

Если теперь на кривой Г в качестве параметра взять переменную длину дуги s: $\Gamma = \{\boldsymbol{r}(s);\ 0 \leqslant s \leqslant S_\Gamma\}$, то, согласно (16.29),

$$\frac{dx}{ds} = \cos\alpha,\ \frac{dy}{ds} = \cos\beta = \sin\alpha,\ \alpha + \beta = \frac{\pi}{2}, \qquad (16.32)$$

где (рис. 84) α — угол, образованный касательной с положительным направлением оси Ox, а β — с положительным направлением оси Oy. Отметим, что эти формулы можно получить применяя к «криволинейному треугольнику» M_0MP (см. рис. 83) формулы, выражающие синус и косинус углов обычного прямоугольного треугольника через его катеты и гипотенузу, считая стороны указанного «треугольника» M_0MP равными соответственно dx, dy, ds. Подобное обстоятельство имеет место и для формулы (16.29), т. е. для случая пространства. Такой метод получения формул (16.29) и (16.32) является, конечно, необоснованным — он не имеет доказательной силы, однако облегчает запоминание этих формул.

16.7. Физический смысл производной векторной функции

Пусть теперь годограф Г непрерывно дифференцируемой векторной функции $\boldsymbol{r}(t)$ есть траектория движущейся материальной точки, а параметр t — время движения. Обозначим переменную длину дуги, отсчитываемую от некоторой начальной точки $\boldsymbol{r}(t_0)$, через $s = s(t)$. Пусть $t > t_0$; положив $\Delta s = s(t + \Delta t) - s(t)$, согласно (16.23), получим

$$\left|\frac{d\boldsymbol{r}}{dt}\right| = \frac{ds}{dt} = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t},$$

т. е. длина вектора $\frac{d\boldsymbol{r}}{dt}$ совпадает с числовым значением скорости в рассматриваемой точке (см. п. 9.4); сам же вектор $\frac{d\boldsymbol{r}}{dt}$, как известно (см. п. 16.2), направлен по касательной.

Вектор $\frac{d\boldsymbol{r}}{dt}$ называется в этом случае *скоростью движения* в данной точке и обозначается $\boldsymbol{v}$: $\boldsymbol{v} = \frac{d\boldsymbol{r}}{dt}$.

§ 17.

Кривизна и кручение кривой

17.1. Две леммы. Радиальная и трансверсальная составляющие скорости

Докажем две полезные для дальнейшего леммы о производных векторных функций.

ЛЕММА 1. *Пусть векторная функция $\boldsymbol{r}(t)$ имеет производную в точке t_0. Если длина вектора $\boldsymbol{r}(t)$ в некоторой окрестности точки t_0 постоянна, то вектор $\boldsymbol{r}'(t_0)$ ортогонален вектору $\boldsymbol{r}(t_0)$, т. е.*

$$\boldsymbol{r}'(t_0)\boldsymbol{r}(t_0) = 0. \tag{17.1}$$

Доказательство. По условию, существует окрестность точки t_0, в которой длина вектора $\boldsymbol{r}(t)$ постоянна: $|\boldsymbol{r}(t)| = c$, где c — константа. Поэтому для всех точек указанной окрестности имеем $|\boldsymbol{r}(t)|^2 = c^2$ $\boldsymbol{r}^2(t) = c^2$. Продифференцировав обе его части в точке t_0, получим (см. п. 15.2) $2\boldsymbol{r}(t_0)\boldsymbol{r}'(t_0) = 0$, откуда и следует (17.1) □.

Утверждение леммы содержательно лишь в случае, когда $\boldsymbol{r}'(t_0) \neq 0$ (если $\boldsymbol{r}'(t_0) = 0$, то условие (17.1) выполняется, очевидно, и без условия постоянства длины вектора $\boldsymbol{r}(t)$). В этом случае физический смысл этой леммы состоит в том, что у материальной точки, движущейся так, что она все время остается на поверхности сферы, ее скорость $\boldsymbol{v} = \boldsymbol{r}' \neq 0$ направлена по касательной к этой сфере и, следовательно, перпендикулярна радиусу-вектору.

Пусть функция $\boldsymbol{r}(t)$ определена в некоторой окрестности $U(t_0)$ точки t_0 и пусть в этой окрестности $\boldsymbol{r}(t) \neq 0$ (если векторная функция $\boldsymbol{r}(t)$ непрерывна в точке t_0, то неравенства нулю радиусов-векторов $\boldsymbol{r}(t)$ в достаточно малой

окрестности точки t_0 всегда можно добиться переносом начала координат). Пусть $t = t_0 + \Delta t \in U(t_0)$ и пусть $\varphi = \varphi(t)$ — угол (выраженный в радианах) между векторами $\boldsymbol{r}(t_0)$ и $\boldsymbol{r}(t)$, $|\varphi| \leqslant \pi$, причем будем считать, что $\varphi(t) \geqslant 0$ для $\Delta t \geqslant 0$ и $\varphi \leqslant 0$ для $\Delta t < 0$; поэтому всегда $\frac{\Delta\varphi}{\Delta t} \geqslant 0$.

Определение 1. *Производная* $\frac{d\varphi(t_0)}{dt}$ *называется угловой скоростью вращения векторной функции* $\boldsymbol{r}(t)$ *в точке* t_0 *и обозначается* $\omega = \omega(t_0; r(t))$:

$$\omega = \frac{d\varphi}{dt}. \qquad (17.2)$$

Заметим, что если выбрать противоположный отсчет углов, т. е. определить угол между векторами $\boldsymbol{r}(t_0)$ и $\boldsymbol{r}(t)$ как угол $\psi = -\varphi$, то очевидно,

$$\frac{d\psi}{dt} \leqslant 0 \text{ и } \omega(t_0; r) = \frac{d\varphi}{dt} = -\frac{d\psi}{dt} = \left|\frac{d\psi}{dt}\right|.$$

Таким образом, как при одном, так и при другом отсчете угла φ между векторами $\boldsymbol{r}(t_0)$ и $\boldsymbol{r}(t)$ всегда

$$\omega(t_0; \boldsymbol{r}) = \left|\frac{d\varphi}{dt}\right|.$$

ЛЕММА 2. *Пусть векторная функция* $\boldsymbol{r}(t)$ *определена в некоторой окрестности точки* t_0 *и* $\boldsymbol{r}(t_0) \neq 0$. *Тогда, если в точке* t_0 *существует производная* $\boldsymbol{r}'(t_0)$, *то в этой точке существует и угловая скорость вращения* $\omega = \omega(t_0; \boldsymbol{r}(t))$, *причем*

$$\omega = \frac{1}{r^2(t_0)}|\boldsymbol{r}(t_0) \times \boldsymbol{r}'(t_0)|. \qquad (17.3)$$

СЛЕДСТВИЕ. *Если в дополнение к условиям леммы длина* r *вектора* $\boldsymbol{r}(t)$ *постоянна*: $|\boldsymbol{r}(t)| = r$ — *константа, то*

$$\omega = \frac{|r'(t)|}{r}. \qquad (17.4)$$

Доказательство. В силу существования производной $\boldsymbol{r}'(t_0)$, функция $\boldsymbol{r}(t)$ непрерывна в точке t_0. Отсюда и из ус-

ловия $\boldsymbol{r}(t_0) \neq 0$ следует, что для всех достаточно малых Δt выполняется неравенство $\boldsymbol{r}(t_0 + \Delta t) \neq 0$ и, следовательно, определен угол $\Delta\varphi$ между векторами $\boldsymbol{r}(t_0)$ и $\boldsymbol{r}(t_0 + \Delta t)$. Из непрерывности векторной функции $\boldsymbol{r}(t)$ в точке t_0 следует также[1] и непрерывность в точке t_0 функции $\varphi(t)$, т. е. $\lim\limits_{\Delta t \to 0} \Delta\varphi = 0$ (как всегда, $\Delta t = t - t_0$, $\Delta\varphi = \varphi(t) - \varphi(t_0) = \varphi(t)$, так как $\varphi(t_0) = 0$).

Для вычисления производной (17.2) заменим бесконечно малую $\Delta\varphi$ эквивалентной ей при $\Delta t \to 0$ бесконечно малой $\sin\Delta\varphi$ (см. лемму в п. 8.2), которую можно найти из формулы

$$|\boldsymbol{r}(t_0) \times \boldsymbol{r}(t_0 + \Delta t)| = |\boldsymbol{r}(t_0)|\,|\boldsymbol{r}(t_0 + \Delta t|\,|\sin\Delta\varphi|.$$

В силу теоремы 2 п. 8.3 о замене бесконечно малых им эквивалентными при вычислении пределов, имеем

$$\omega = \frac{d\varphi}{dt} = \lim_{\Delta t \to 0} \frac{\Delta\varphi}{\Delta t} = \lim_{\Delta t \to 0} \left|\frac{\Delta\varphi}{\Delta t}\right| =$$
$$= \lim_{\Delta t \to 0} \left|\frac{\sin\Delta\varphi}{\Delta t}\right| = \lim_{\Delta t \to 0} \frac{|\boldsymbol{r}(t_0) \times \boldsymbol{r}(t_0 + \Delta t)|}{|\boldsymbol{r}(t_0)||\boldsymbol{r}(t_0 + \Delta t)||\Delta t|} =$$
$$= \frac{1}{\boldsymbol{r}^2(t_0)} \lim_{\Delta t \to 0} \frac{|\boldsymbol{r}(t_0) \times \boldsymbol{r}(t_0 + \Delta t)|}{|\Delta t|}. \qquad (17.5)$$

Здесь снова была использована непрерывность векторной функции $\boldsymbol{r}(t)$ в точке t_0: $\lim\limits_{\Delta t \to 0} \boldsymbol{r}(t_0 + \Delta t) = \boldsymbol{r}(t_0)$.

Далее, поскольку функция $\boldsymbol{r}(t)$ дифференцируема в точке t_0,

$$\boldsymbol{r}(t_0 + \Delta t) = \boldsymbol{r}(t_0) + \boldsymbol{r}'(t_0)\Delta t + \varepsilon(\Delta t)\Delta t,$$

где $\lim\limits_{\Delta t \to 0} \boldsymbol{\varepsilon}(\Delta t) = 0$. Подставив это выражение для $\boldsymbol{r}(t_0 + \Delta t)$ в (17.5) и заметив, что $|\boldsymbol{r}(t_0) \times \boldsymbol{r}(t_0)| = 0$, $\lim\limits_{\Delta t \to 0} |\boldsymbol{r}(t_0) \times \varepsilon(\Delta t)| = 0$, получим

$$\omega = \frac{d\varphi}{dt} = \lim_{\Delta t \to 0} \frac{\Delta\varphi}{\Delta t} = \frac{|\boldsymbol{r}(t_0) \times \boldsymbol{r}'(t_0)|}{\boldsymbol{r}^2(t_0)}. \ \square$$

---

[1] Это вытекает, например, из равенства $\cos\varphi = \dfrac{\boldsymbol{r}(t_0)\boldsymbol{r}(t)}{|\boldsymbol{r}(t_0)\boldsymbol{r}(t)|}$.

Доказательство следствия. Если $|\boldsymbol{r}(t)| = r$ — постоянная, то, в силу леммы 1, $\boldsymbol{r}(t_0)\, \boldsymbol{r}'(t_0) = 0$, т. е. $|\boldsymbol{r}(t_0)|\, |\boldsymbol{r}'(t_0)| \cos \widehat{\boldsymbol{r}\boldsymbol{r}'} = 0$. Так как по условию леммы $|\boldsymbol{r}(t_0)| \neq 0$, то либо $|\boldsymbol{r}'(t_0)| = 0$, либо угол $\widehat{\boldsymbol{r}\boldsymbol{r}'}$ между векторами $\boldsymbol{r}(t_0)$ и $\boldsymbol{r}'(t_0)$ равен $\pm\pi/2$ и, следовательно, $|\sin \widehat{\boldsymbol{r}\boldsymbol{r}'}| = 1$. В обоих случаях

$$|\boldsymbol{r}(t_0) \times \boldsymbol{r}'(t_0)| = |\boldsymbol{r}(t_0)|\, |\boldsymbol{r}'(t_0)|\, |\sin \widehat{\boldsymbol{r}\boldsymbol{r}'}| = r|\boldsymbol{r}'(t_0)|.$$

Подставляя это выражение в формулу (17.3), получим формулу (17.4). □

Леммы 1 и 2 остаются справедливыми и в том случае, если под окрестностями понимать односторонние окрестности.

Для выяснения физического смысла формул (17.3) и (17.4) будем снова интерпретировать кривую, описываемую концом радиуса-вектора $\boldsymbol{r}(t)$, как траекторию движения материальной точки, а параметр t — как время. Пусть длина вектора $\boldsymbol{r}(t)$ остается постоянной: $|\boldsymbol{r}(t)| = r$, т. е. точка движется по сфере радиуса r. Рассмотрим движение точки в каждый момент времени как вращение около так называемой мгновенной оси вращения, т. е. оси, проходящей через начало координат, перпендикулярно плоскости движения $\Big($так называется плоскость, проходящая через радиус-вектор $\boldsymbol{r}(t)$ параллельно скорости $\boldsymbol{v} = \frac{d\boldsymbol{r}(t)}{dt}\Big)$. Тогда вектор $\boldsymbol{\omega} = \frac{\boldsymbol{r} \times \boldsymbol{r}'}{\boldsymbol{r}^2}$ физически означает вектор угловой скорости, а формулы (17.3) и (17.4) выражают связь между угловой скоростью ω и линейной скоростью $\boldsymbol{v}$. В частности, формула (17.4) в этих обозначениях принимает вид

$$|\boldsymbol{\omega}| = \frac{|\boldsymbol{v}|}{r}.$$

Замечание. Используя лемму 1, можно легко получить полезное разложение производной векторной функции на две ортогональные составляющие: в направлении вектора $\boldsymbol{r}(t)$ (*радиальная составляющая*) и в перпендикулярном направлении (*трансверсальная составляющая*).

Пусть векторная функция $\boldsymbol{r}(t)$ определена в некоторой окрестности точки t_0, $\boldsymbol{r}(t_0) \neq 0$, и существует производная $\boldsymbol{r}'(t_0)$. Положим $\boldsymbol{r}_0(t) = \frac{\boldsymbol{r}(t)}{|\boldsymbol{r}(t)|}$, очевидно, $|\boldsymbol{r}_0(t)| \equiv 1$. В точке t_0 существует производная

$$\frac{d|\boldsymbol{r}|}{dt} = \frac{d}{dt}\sqrt{\boldsymbol{r}^2} = \frac{\boldsymbol{r}\boldsymbol{r}'}{|\boldsymbol{r}|} = \boldsymbol{r}_0\boldsymbol{r}',$$

следовательно, в точке t_0 существует и производная $\frac{d\boldsymbol{r}_0}{dt}$, которая, согласно лемме 1, ортогональна вектору $\boldsymbol{r}_0(t_0)$, а поэтому и вектору $\boldsymbol{r}(t_0)$.

Дифференцируя равенство $\boldsymbol{r}(t) = |\boldsymbol{r}(t)|\boldsymbol{r}_0(t)$ в точке t_0, получим

$$\frac{d\boldsymbol{r}}{dt} = \frac{d|\boldsymbol{r}|}{dt}\boldsymbol{r}_0 + |\boldsymbol{r}|\frac{d\boldsymbol{r}_0}{dt} = (\boldsymbol{r}_0\boldsymbol{r}')\boldsymbol{r}_0 + |\boldsymbol{r}|\frac{d\boldsymbol{r}_0}{dt}.$$

Это и есть искомое разложение.

В том случае, если годограф векторной функции $\boldsymbol{r}(t)$ является траекторией движущейся материальной точки, полученная формула дает разложение ее скорости на составляющую поступательного движения (радиальная составляющая) и составляющую вращательного движения (трансверсальная составляющая).

17.2. Определение кривизны кривой и ее вычисление

Рассмотрим дважды дифференцируемую кривую Γ без особых точек. Такая кривая спрямляема и у нее существует дважды дифференцируемое представление $\boldsymbol{r} = \boldsymbol{r}(s)$, в котором за параметр принята переменная длина дуги s. Пусть $0 \leqslant s_0 \leqslant S$, $\Delta s = s - s_0$, а $\alpha = \alpha(s)$ — угол между касательными к кривой Γ в точках $r(s_0)$ и $r(s_0 + \Delta s)$, причем будем считать, что $\alpha(s) \geqslant 0$ для $\Delta s \geqslant 0$, $\alpha(s) \leqslant 0$ для $\Delta s < 0$ и $|\alpha| \leqslant \pi/2$. Очевидно, $\Delta\alpha = \alpha(s) - \alpha(s_0) = \alpha(s)$, так как $\alpha(s_0) = 0$.

Пусть теперь $\boldsymbol{t}(s) = \frac{d\boldsymbol{r}(s)}{ds}$. Как было показано, $\boldsymbol{t}(s)$ является единичным вектором (см. (16.26)), параллельным касательной к кривой в соответствующей точке (см. п. 16.4), поэтому угол $\Delta\alpha$ является и углом между векторами $\boldsymbol{t}(s_0)$ и $\boldsymbol{t}(s_0 + \Delta s)$.

Определение 2. *Угловая скорость вращения касательного единичного вектора* $\boldsymbol{t} = \frac{d\boldsymbol{r}}{ds}$ *в данной точке кривой называется кривизной* $k(s_0)$ *кривой в этой точке*

$$k(s_0) = \omega(s_0;\, t) = \frac{d\alpha(s_0)}{ds}.$$

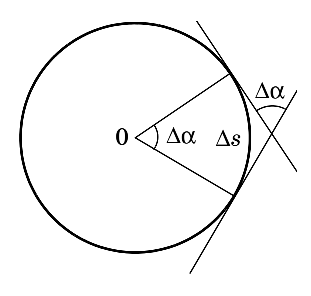


Рис. 85

Опуская для краткости значение аргумента, получаем

$$k \overset{\text{def}}{=} \frac{d\alpha}{ds}. \qquad (17.6)$$

Кривая Г дважды дифференцируема, поэтому существует производная $\frac{d\boldsymbol{t}}{ds} = \frac{d^2\boldsymbol{r}}{ds^2}$ и так как вектор $\boldsymbol{t}$ единичный, то, в силу следствия леммы 2 из п. 17.1, отсюда имеем

$$k = \frac{d\alpha}{ds} = \left|\frac{d^2 r}{ds^2}\right| = \left|\frac{dt}{ds}\right|, \text{ ибо } \boldsymbol{t} = \frac{dr}{ds}. \qquad (17.7)$$

Определение 3. *Величина, обратная кривизне, называется радиусом кривизны в данной точке и обозначается* R, *т. е.* $R = 1/k$.

Пусть Г — окружность радиуса R. В этом случае угол $\Delta\alpha$ между касательными равен углу, образованному радиусами, проведенными в точки касания (рис. 85), а для длины дуги Δs между этими точками имеет место формула $\Delta s = R\Delta\alpha$. Поэтому $\left|\frac{\Delta\alpha}{\Delta s}\right| = \frac{1}{R}$. По определению же кривизны для окружности имеем

$$k = \lim_{\Delta s \to 0}\left|\frac{\Delta\alpha}{\Delta s}\right| = \frac{1}{R}.$$

Таким образом, в случае окружности ее кривизна k постоянна (не зависит от точки) и равна обратной величине радиуса; радиус же кривизны окружности равен ее радиусу. Отсюда и произошел термин «радиус кривизны».

Обозначим через $\boldsymbol{n}$ единичный вектор в направлении вектора $\frac{dt}{ds}$. Из формулы (17.7) следует, что вектор $\boldsymbol{n}$ однозначно

определен лишь для тех точек, в которых кривизна k не равна нулю, и что в этих точках

$$\frac{d\boldsymbol{t}}{ds} = k\boldsymbol{n}. \qquad (17.8)$$

Вектор $\boldsymbol{t}$ — единичный, поэтому его производная, а следовательно, и вектор $\boldsymbol{n}$ перпендикулярны ему (см. лемму 1 в п. 17.1): $\boldsymbol{n} \perp \boldsymbol{t}$.

Определение 4. *Вектор $\boldsymbol{n}$ называется вектором главной нормали (короче, главной нормалью) кривой Γ в данной ее точке.*

Достаточные условия существования кривизны в данной точке и метод ее вычисления дает следующая теорема.

ТЕОРЕМА 1. *Пусть $\Gamma = \{\boldsymbol{r}(t);\ a \leqslant t \leqslant b\}$ — дважды дифференцируемая кривая без особых точек. Тогда в каждой ее точке существует кривизна k и она следующим образом выражается через производную по переменной длине дуги:*

$$k = \left|\frac{d^2\boldsymbol{r}}{ds^2}\right| \qquad (17.9)$$

и через производные по произвольному параметру:

$$k = \frac{|\boldsymbol{r}' \times \boldsymbol{r}''|}{|\boldsymbol{r}'|^3}. \qquad (17.10)$$

Штрих здесь и в дальнейшем обозначает производные по произвольному параметру t. Производные по длине дуги s будем обозначать символом $\frac{d}{ds}$.

Доказательство. При предположениях теоремы переменная длина дуги $s = s(t)$, $a \leqslant t \leqslant b$, $0 \leqslant s \leqslant S$, кривой Γ может быть принята на этой кривой за параметр (см. следствие 2 из теоремы 2 в п. 16.5). При этом единичный касательный вектор $\boldsymbol{t} = \frac{d\boldsymbol{r}}{ds}$ является дифференцируемой векторной функцией, поэтому у него существует производная $\frac{d\boldsymbol{t}}{ds} = \frac{d^2\boldsymbol{r}}{ds^2}$, следовательно, и кривизна $k = \left|\frac{d\boldsymbol{t}}{ds}\right| = \left|\frac{d^2\boldsymbol{r}}{ds^2}\right|$ (см. (17.7)). Формула (17.9) доказана.

Для того чтобы доказать формулу (17.10), рассмотрим длину векторного произведения векторов

$$\frac{d^2\boldsymbol{r}}{ds^2} = \frac{d\boldsymbol{t}}{ds} \underset{(17.8)}{=} k\boldsymbol{n} \text{ и } \frac{d\boldsymbol{r}}{ds} = \boldsymbol{t}.$$

Заметив, что $k \geqslant 0$, что вектор $\boldsymbol{t}$ перпендикулярен вектору $\boldsymbol{n}$, а длины их равны единице, получим

$$\left|\frac{d^2\boldsymbol{r}}{ds^2} \times \frac{d\boldsymbol{r}}{ds}\right| = k|\boldsymbol{n} \times \boldsymbol{t}| = k. \tag{17.11}$$

Выразив производные $\frac{d^2\boldsymbol{r}}{ds^2}$ и $\frac{d\boldsymbol{r}}{ds}$ через производные по t:

$$\frac{d\boldsymbol{r}}{ds} = \boldsymbol{r}'\frac{dt}{ds} = \frac{\boldsymbol{r}'}{s'},$$

$$\frac{d^2\boldsymbol{r}}{ds^2} = \frac{d}{ds}\left(\frac{\boldsymbol{r}'}{s'}\right) = \frac{s'\boldsymbol{r}'' - s''\boldsymbol{r}'}{s'^2}\frac{dt}{ds} = \frac{s'\boldsymbol{r}'' - s''\boldsymbol{r}'}{s'^3}, \tag{17.12}$$

будем иметь

$$\left|\frac{d^2\boldsymbol{r}}{ds^2} \times \frac{d\boldsymbol{r}}{ds}\right| = \left|\frac{s'\boldsymbol{r}'' - s''\boldsymbol{r}'}{s'^3} \times \frac{\boldsymbol{r}'}{s'}\right| = \left|\frac{\boldsymbol{r}' \times \boldsymbol{r}''}{s'^3}\right| = \frac{|\boldsymbol{r}' \times \boldsymbol{r}''|}{|\boldsymbol{r}'|^3}, \tag{17.13}$$

поскольку $\boldsymbol{r}' \times \boldsymbol{r}' = \boldsymbol{0}$ и $|s'| \underset{(16.13)}{=} |\boldsymbol{r}'|$.

В равенствах (17.11) и (17.13) равны левые части, поэтому равны правые части, а это означает справедливость формулы (17.10). □

От формулы (17.10) легко перейти к выражению для кривизны в координатной записи. В самом деле, замечая, что $\boldsymbol{r}' = (x', y', z')$, $\boldsymbol{r}'' = (x'', y'', z'')$ и что

$$\boldsymbol{r}' \times \boldsymbol{r}'' = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ x' & y' & z' \\ x'' & y'' & z'' \end{vmatrix}$$

(где $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ — единичные векторы соответственно в направлении осей Ox, Oy, Oz), получаем

$$|\boldsymbol{r}' \times \boldsymbol{r}''| =$$

$$= \sqrt{(y'z'' - y''z')^2 + (z'x'' - z''x')^2 + (x'y'' - x''y')^2}, \tag{17.14}$$

с другой стороны,

$$|\boldsymbol{r}'| = \sqrt{x'^2 + y'^2 + z'^2}. \tag{17.15}$$

Подставив (17.14) и (17.15) в (17.10), мы и найдем искомое выражение.

17.3. Главная нормаль. Соприкасающаяся плоскость

Рассмотрим некоторые свойства вектора главной нормали $\boldsymbol{n}$.

Вектор $\frac{d^2\boldsymbol{r}}{ds^2}$, поэтому и вектор $\boldsymbol{n} = \frac{1}{k}\frac{d^2\boldsymbol{r}}{ds^2}$ не зависят от выбора ориентации кривой. Действительно, если σ — переменная длина дуги кривой, отсчитываемая в противоположном направлении, и, следовательно, если $\sigma = S - s$, то, заметив, что $\frac{d\sigma}{ds} = -1$, получим

$$\frac{d^2r}{ds^2} = \frac{d^2r}{d\sigma^2}\left(\frac{d\sigma}{ds}\right)^2 = \frac{d^2\boldsymbol{r}}{d\sigma^2}.$$

Определение 5. *Всякая прямая, проходящая через точку кривой и перпендикулярная касательной в этой точке, называется нормалью к кривой в данной точке. Нормаль к кривой, параллельная вектору n, называется главной нормалью.*

Вектор главной нормали $\boldsymbol{n}$ при $\Delta s \to 0$ с точностью до бесконечно малых более высокого порядка, чем Δs^2, указывает направление, в котором кривая в окрестности данной точки отклоняется от своей касательной (рис. 86). Действительно, выбрав на кривой в качестве параметра переменную длину дуги s, согласно формуле Тейлора для векторной функции (см. п. 15.2), будем иметь

$$\Delta\boldsymbol{r} = \boldsymbol{r}(s_0 + \Delta s) - \boldsymbol{r}(s_0) =$$

$$= \frac{dr(s_0)}{ds}\Delta s + \frac{1}{2}\frac{d^2\boldsymbol{r}(s_0)}{ds^2}\Delta s^2 + o(\Delta s^2),\ \ \Delta s \to 0, \qquad (17.16)$$

или, заметив, что (см. 17.8))

$$\frac{d\boldsymbol{r}}{ds} = t,\ \ \frac{d^2\boldsymbol{r}}{ds^2} = \frac{dt}{ds} = k\boldsymbol{n}, \qquad (17.17)$$

получим

$$\Delta\boldsymbol{r} - \Delta s\boldsymbol{t} = \frac{1}{2}k\Delta s^2\boldsymbol{n} + \boldsymbol{o}(\Delta s^2),\ \ \Delta s \to 0;$$

поскольку $\frac{1}{2}k\Delta s^2 > 0$, эта формула и доказывает справедливость нашего утверждения.

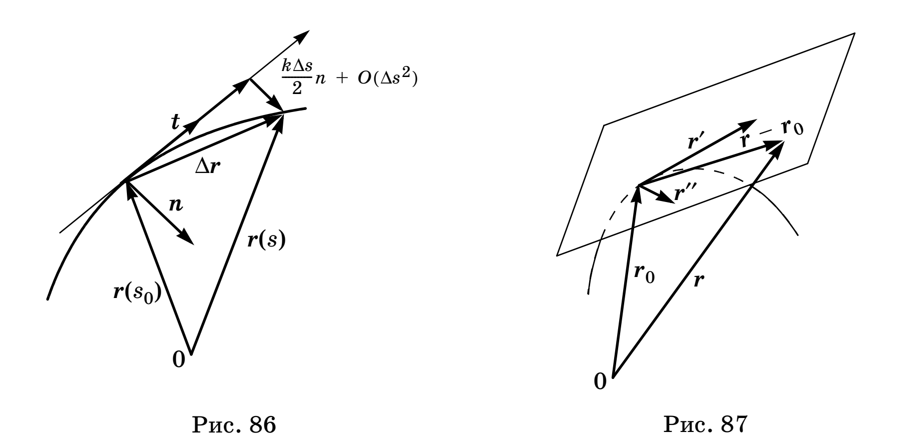


Рис. 86

Рис. 87

Определение 6. *Плоскость, проходящая через касательную и главную нормаль в данной точке кривой, называется соприкасающейся плоскостью.*

В силу этого определения, соприкасающаяся плоскость определена для точек, в которых $k \neq 0$.

Эта плоскость обладает тем свойством, что дважды непрерывно дифференцируемая кривая в окрестности каждой своей точки, в которой кривизна не равна нулю, лежит «почти» в соприкасающейся плоскости. Это означает, что конец радиуса-вектора $\boldsymbol{r}(s_0 + \Delta s)$ отстоит от конца радиуса-вектора $\boldsymbol{r}(s_0) + \Delta s\boldsymbol{t} + \frac{1}{2}k\Delta s^2\boldsymbol{n}$, лежащего, очевидно, на соприкасающейся плоскости, на величину бесконечно малую по сравнению с Δs^2. Это сразу следует из равенства (17.16):

$$\boldsymbol{r}(s_0 + \Delta s) - \left(\boldsymbol{r}(s_0) + \Delta s\boldsymbol{t} + \frac{1}{2}k\Delta s^2\boldsymbol{n}\right) = \boldsymbol{o}(\Delta s^2).$$

Найдем уравнение соприкасающейся плоскости для кривой, заданной представлением $\boldsymbol{r} = \boldsymbol{r}(t)$ с произвольным параметром t. Как и выше, производные по переменному t будем обозначать штрихом, а производные по длине дуги s — символом $\frac{d}{ds}$. Дифференцируя $\boldsymbol{r} = \boldsymbol{r}(t)$ как сложную функцию $\boldsymbol{r} = \boldsymbol{r}(s)$, $s = s(t)$, получим (см. (17.17))

$$\boldsymbol{r}' = \frac{d\boldsymbol{r}}{ds}s' = s'\boldsymbol{t},$$

$$\boldsymbol{r}'' = s'^2\frac{d\boldsymbol{t}}{ds} + s''\boldsymbol{t} = s'^2k\boldsymbol{n} + s''\boldsymbol{t}. \qquad (17.18)$$

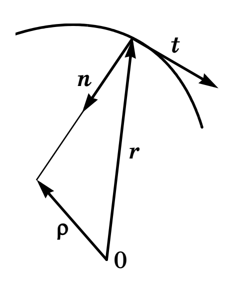


Рис. 88

Эти формулы, очевидно, являются обращением формул (17.12).

Отсюда следует, что векторы $\boldsymbol{r}'$ и $\boldsymbol{r}''$ также параллельны соприкасающейся плоскости; в силу же условия $k \neq 0$ выполняется неравенство $\boldsymbol{r}' \times \boldsymbol{r}'' \neq 0$ (см. (17.10)) и, следовательно, $\boldsymbol{r}'$ и $\boldsymbol{r}''$ не коллинеарны. Обозначим теперь через $\boldsymbol{r}_0$, $\boldsymbol{r}_0'$ и $\boldsymbol{r}_0''$ векторы $\boldsymbol{r}$, $\boldsymbol{r}'$ и $\boldsymbol{r}''$ в некоторой фиксированной точке данной кривой Г, а через $\boldsymbol{r}$ — текущий вектор соприкасающейся плоскости; тогда смешанное произведение векторов $\boldsymbol{r} - \boldsymbol{r}_0$, $\boldsymbol{r}_0'$ и $\boldsymbol{r}_0''$ должно быть равно нулю, так как все они параллельны соприкасающейся плоскости (рис. 87):

$$(\boldsymbol{r} - \boldsymbol{r}_0, \boldsymbol{r}_0', \boldsymbol{r}_0'') = 0.$$

Это и есть уравнение указанной плоскости в векторном виде. В координатном виде оно запишется следующим образом:

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ x' & y_0' & z_0' \\ x_0'' & y_0'' & z_0'' \end{vmatrix} = 0,$$

где $\boldsymbol{r} = (x, y, z)$, $\boldsymbol{r}_0 = (x_0, y_0, z_0)$, $\boldsymbol{r}_0' = (x_0', y_0', z_0')$, $\boldsymbol{r}_0'' = (x_0'', y_0'', z_0'')$.

17.4. Центр кривизны и эволюта кривой

Определение 7. *Точка пространства, лежащая на главной нормали, проведенной в данной точке кривой, и находящаяся от этой точки кривой на расстоянии, равном радиусу кривизны R, в направлении вектора главной нормали $\boldsymbol{n}$, называется центром кривизны кривой в указанной ее точке* (рис. 88).

Таким образом, если ρ — радиус-вектор центра кривизны, а $\boldsymbol{r}$, как обычно, радиус-вектор данной точки кривой, то

$$\rho = \boldsymbol{r} + R\boldsymbol{n},$$

или, так как $R = \frac{1}{k}$, а $\boldsymbol{n} \underset{(17.8)}{=} \frac{1}{k}\left|\frac{dt}{ds}\right| = \frac{1}{k}\frac{d^2\boldsymbol{r}}{ds^2}$, то

$$\rho = \boldsymbol{r} + \frac{1}{k^2}\frac{d^2\boldsymbol{r}}{ds^2}. \tag{17.19}$$

Найдем выражение для ρ через производные векторной функции по произвольному параметру t. Для этого подставим в получившуюся формулу выражение для второй производной $\frac{d^2\boldsymbol{r}}{ds^2}$ через производные по t (см. (17.12)), в результате получим

$$\rho = \boldsymbol{r} + \frac{1}{k^2}\frac{s'\boldsymbol{r}'' - s''\boldsymbol{r}'}{s'^3}, \qquad (17.20)$$

где (считая для простоты, что при возрастании параметра t длина дуги $s(t)$ также возрастает) $s' = |\boldsymbol{r}'| = \sqrt{x'^2 + y'^2 + z'^2}$, откуда

$$s'' = \frac{x'y'' + y'x'' + z'z''}{\sqrt{x'^2 + y'^2 + z'^2}}.$$

Формулу (17.19) можно рассматривать как представления некоторой кривой, точками носителя которой являются центры кривизны в точках данной кривой. Эта кривая называется *эволютой* данной кривой.

17.5. Формулы для кривизны и эволюты плоской кривой

Все сказанное в предыдущем пункте, в частности, справедливо и для плоских кривых. Заметим лишь, что если кривая $\Gamma = \{\boldsymbol{r}(t)\}$ лежит в некоторой плоскости, то все производные векторной функции $\boldsymbol{r}(t)$, если их начало поместить на эту плоскость, будут также лежать в ней. В самом деле, в ней лежит приращение векторной функции $\Delta\boldsymbol{r} = \boldsymbol{r}(t + \Delta t) - \boldsymbol{r}(t)$, поэтому и отношение $\frac{\Delta\boldsymbol{r}}{\Delta t}$. Отсюда легко следует, что и предел этих отношений $\boldsymbol{r}' = \lim\limits_{\Delta t \to 0}\frac{\Delta\boldsymbol{r}}{\Delta t}$ лежит в указанной плоскости. Применяя то же рассуждение к $\boldsymbol{r}'$, мы докажем, что и $\boldsymbol{r}''$ находится в той же плоскости, и т. д.

Из сказанного следует, что если кривая лежит в некоторой плоскости, то касательный вектор $\boldsymbol{t}'$, а если ее кривизна $k \neq 0$, то и вектор главной нормали $\boldsymbol{n}$ лежат в той же плоскости. Поэтому эта плоскость является соприкасающейся плоскостью для рассматриваемой кривой.

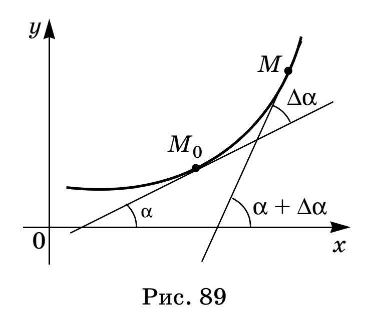


Рис. 89

Отметим также, что если в случае кривой $\Gamma = \{\boldsymbol{r}(s);\ 0 \leqslant s \leqslant S\}$, лежащей в плоскости xOy, в отличие от п. 17.2 через $\alpha(s)$ обозначить угол, образованный касательной в точке $r(s)$ с осью Ox (рис. 89), то $\Delta\alpha = \alpha(s_0 + \Delta s) - \alpha(s_0)$ является углом между касательными в точках $r\ (s_0)$ и $r\ (s_0 + \Delta s)$, однако его знак может оказаться другим, чем в п. 17.2.

Если угол α возрастает вместе с s, т. е. если $\frac{\Delta\alpha}{\Delta s} \geqslant 0$ при $\Delta s > 0$, то

$$k = \lim_{\Delta s \to 0} \frac{\Delta\alpha}{\Delta s} = \frac{d\alpha}{ds};$$

если же α убывает с возрастанием s, то

$$k = -\lim_{\Delta s \to 0} \frac{\Delta\alpha}{\Delta s} = -\frac{d\alpha}{ds}.$$

Запишем некоторые из формул, полученных в предыдущем пункте, считая, что кривая $\Gamma = \{\boldsymbol{r}(t);\ a \leqslant t \leqslant b\}$ лежит в плоскости xOy: $\boldsymbol{r}(t) = (x(t), y(t))$. Из формул (17.10), (17.14) и (17.15) имеем

$$k = \frac{1}{R} = \frac{|x'y'' - x''y'|}{(x'^2 + y'^2)^{3/2}}. \qquad (17.21)$$

Обозначая через (ξ, η) центр кривизны кривой Γ, из формулы (17.19) получим формулы, выражающие координаты ξ и η через производные по s:

$$\xi = x + R^2\frac{d^2x}{ds^2},\ \ \eta = y + R^2\frac{d^2y}{ds^2},$$

а из формулы (17.20) следуют формулы, выражающие координаты центра кривизны через производные по произвольному параметру t:

$$\xi = x + \frac{(x'^2 + y'^2)^3}{(x'y'' - x''y')^2}\,\frac{x''\sqrt{x'^2 + y'^2} - x'\dfrac{x'x'' + y'y''}{\sqrt{x'^2 + y'^2}}}{(x'^2 + y'^2)^{3/2}} =$$

$$= x - y'\frac{x'^2 + y'^2}{x'y'' - x''y'}; \qquad (17.22)$$

аналогично,

$$\eta = y + x'\frac{x'^2 + y'^2}{x'y'' - x''y'}. \qquad (17.23)$$

УПРАЖНЕНИЕ 1. Пусть Γ — дважды дифференцируемая плоская кривая без особых точек, пусть α — угол наклона ее касательной к оси Ox и пусть $k^* = \frac{d\alpha}{ds}$ (следовательно, $|k^*| = k$) и $R^* = \frac{1}{k^*}$. Показать, что $\xi = x - R^*\sin\alpha$, $\eta = y + R^*\cos\alpha$, а также что $\xi = x - \frac{dy}{d\alpha}$, $\eta = y + \frac{dx}{d\alpha}$.

В том случае, когда кривая является графиком функции $y = f(x)$, формулы (17.21), (17.22) и (17.23) принимают вид

$$k = \frac{|y''|}{(1 + y'^2)^{3/2}}, \qquad (17.24)$$

$$\xi = x - \frac{1 + y'^2}{y''}\, y', \quad \eta = y + \frac{1 + y'^2}{y''}. \qquad (17.25)$$

Примеры. 1. Найдем кривизну и эволюту параболы $y = ax^2$, $a > 0$.

Замечая, что $y' = 2ax$, $y'' = 2a$, по формуле (17.24) имеем $k = \dfrac{2a}{(1 + 4a^2x^2)^{3/2}}$. Чтобы найти уравнение эволюты, воспользуемся формулами (17.25):

$$\xi = x - \frac{1 + 4a^2x^2}{2a}\,2ax = -4a^2x^3,$$

$$\eta = ax^2 + \frac{1 + 4a^2x^2}{2a} = \frac{6a^2x^2 + 1}{2a}.$$

Получено параметрическое представление эволюты параболы с параметром x. Можно получить и ее явное представление, исключив этот параметр x. Для этого из первого равенства найдем $x^3 = -\dfrac{\xi}{4a^2}$, а из второго $x^2 = \dfrac{2a\eta - 1}{6a^2}$. Возводя первое из равенств в квадрат, а второе — в куб и приравнивая правые части, имеем

$$\left(\frac{\xi}{4a^2}\right)^2 = \left(\frac{2a\eta - 1}{6a^2}\right)^3,$$

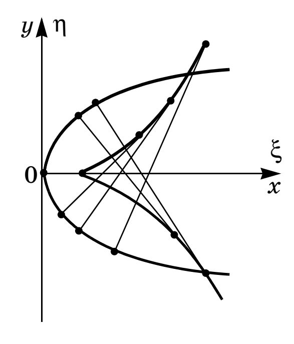


Рис. 90

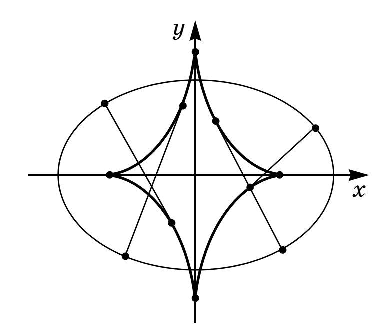


Рис. 91

откуда

$$\xi = \pm\frac{4}{3}\sqrt{\frac{a}{3}}\left(\eta - \frac{1}{2a}\right)^{3/2}.$$

Эта кривая, изображенная на рисунке 90, является, как известно (см. пример 2 в п. 14.3), полукубической параболой.

2. Найдем радиус кривизны и эволюту эллипса $x = a\cos t$, $y = b\sin t$, $a \geqslant b > 0$.

Заметив, что $x' = -a\sin t$, $y' = b\cos t$, $x'' = -a\cos t$, $y'' = -b\sin t$, по формуле (17.21) получим

$$R = \frac{1}{k} = \frac{(a^2\sin^2 t + b^2\cos^2 t)^{3/2}}{ab\sin^2 t + ab\cos^2 t} = \frac{(a^2\sin^2 t + b^2\cos^2 t)^{3/2}}{ab}.$$

Поэтому из формул (17.22) и (17.23) следует, что

$$\xi = a\cos t - b\cos t\,\frac{a^2\sin^2 t + b^2\cos^2 t}{ab} = \frac{a^2 - b^2}{a}\cos^3 t,$$

$$\eta = b\sin t - a\sin t\,\frac{a^2\sin^2 t + b^2\cos^2 t}{ab} = \frac{b^2 - a^2}{b}\sin^3 t$$

— параметрическое представление искомой эволюты; параметр t можно исключить, возводя получившиеся равенства в степень 2/3 и складывая их:

$$(a\xi)^{2/3} + (b\eta)^{2/3} = (a^2 - b^2)^{2/3}.$$

Эта кривая называется *астроидой* (рис. 91).

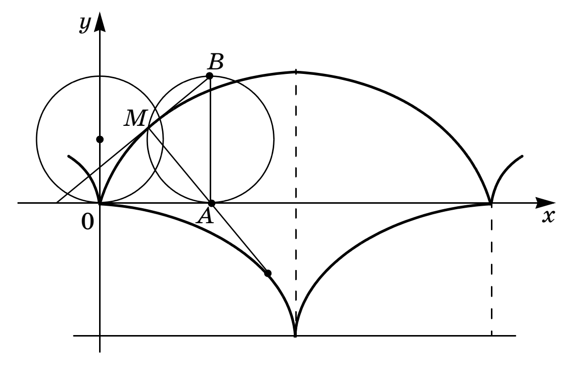


Рис. 92

3. Найдем радиус кривизны и эволюту циклоиды (см. примеры в п. 16.2 и 16.4):

$$x = r\,(t - \sin t),\ y = r\,(1 - \cos t),\ 0 \leqslant t \leqslant 2\pi.$$

Если $s = s(t)$ — переменная длина дуги циклоиды, то

$$s' = \sqrt{x'^2 + y'^2} = r\sqrt{(1 - \cos t)^2 + \sin^2 t} = 2r \sin \frac{t}{2}, \quad (17.26)$$

а так как для угла α, образованного касательной к циклоиде с осью абсцисс, имеет место формула $\alpha = \pi/2 - t/2$ (см. п. 16.4), то $\alpha' = -1/2$, поэтому для радиуса кривизны циклоиды имеем

$$R = \frac{1}{k} = \left|\frac{ds}{d\alpha}\right| = \left|\frac{s'}{\alpha'}\right| = 4r\sin \frac{t}{2}. \qquad (17.27)$$

Если соединить точку циклоиды M с нижней точкой A катящейся окружности и обозначить через B верхнюю точку этой окружности (рис. 92), то угол AMB, как опирающийся на диаметр окружности, будет прямым. Поскольку прямая MB является касательной (см. п. 6.4), MA представляет собой нормаль к циклоиде. При этом $|MA| = 2r \sin \frac{t}{2}$, так как $\angle MAB = \alpha = \pi/2 - t/2$. Следовательно, $R = 2|MA|$.

Отсюда следует простой способ нахождения центра кривизны циклоиды. Для этого надо соединить точку M циклоиды

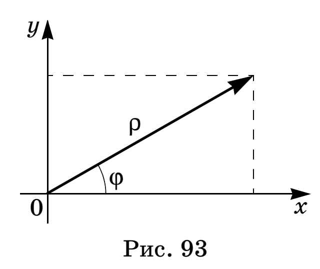


Рис. 93

с нижней точкой A катящейся окружности и на получившейся прямой отложить от точки M в сторону точки A отрезок в два раза большей длины, чем длина отрезка MA. Конец отложенного отрезка и будет центром кривизны циклоиды.

Найдем его координаты ξ, η по формулам (17.22) и (17.23). Имеем:

$$x'y'' - x''y' = r^2((1 - \cos t)\cos t - \sin^2 t) =$$

$$= r^2(\cos t - 1) = -2r^2\sin^2 \frac{t}{2}.$$

Поэтому

$$\xi \underset{\substack{(17.22)\\(17.26)}}{=} x + 2y' = x + 2r\sin t = r(t + \sin t),$$

$$\eta \underset{\substack{(17.23)\\(17.26)}}{=} y - 2x' = -r(1 - \cos t).$$

Введем новый параметр u по формуле $t = u + \pi$. Тогда получим

$$\xi = r(u - \sin u) + \pi r, \quad \eta = r(1 - \cos u) - 2r, \quad -\pi \leqslant u \leqslant \pi.$$

Таким образом, эволютой циклоиды является та же самая циклоида, только параллельно перенесенная (см. рис. 92).

Иногда для изображения кривой бывает удобно использовать так называемые *полярные координаты* (ρ, φ), $\rho \geqslant 0$, $-\pi < \varphi \leqslant \pi$, где ρ — длина радиуса-вектора данной точки M, а φ — угол, образованный радиусом-вектором с осью Ox, $-\pi < \varphi \leqslant \pi$. Таким образом, каждой точке плоскости, кроме начала координат, взаимно однозначно соответствует указанная упорядоченная пара (ρ, φ), $\rho > 0$; для начала же координат имеем $\rho = 0$, а угол φ не определен (рис. 93).

Если $M = (x, y)$, где, как обычно, x и y — декартовы координаты точки M, то

$$x = \rho \cos \varphi, \quad y = \rho \sin \varphi. \tag{17.28}$$

Обратная связь выражается формулами

$$\rho = \sqrt{x^2 + y^2}, \quad \varphi = \operatorname{arctg} \frac{y}{x} + k\pi,$$

где $k = 0$, если $x \geqslant 0$; $k = 1$, если $x < 0$, $y > 0$, и $k = -1$, если $x < 0$, $y < 0$; при этом, как обычно, при $x = 0$, $y \neq 0$ считается $\operatorname{arctg} \frac{y}{x} = \frac{\pi}{2} \operatorname{sign} y$.

Иногда на угол φ не накладывают ограничения $-\pi < \varphi \leqslant \pi$, а обозначают через φ любой угол, для которого $\operatorname{tg} \varphi = \frac{y}{x}$. В этом случае соответствие между упорядоченными парами (ρ, φ), $\rho \neq 0$, и точками плоскости, отличными от начала координат, уже, очевидно, не является взаимно однозначным.

Если задана непрерывная функция

$$\rho = \rho(\varphi), \ \alpha \leqslant \varphi \leqslant \beta, \tag{17.29}$$

то, подставляя ее в (17.28), получаем

$$x = \rho(\varphi)\cos \varphi, \ y = \rho(\varphi)\sin \varphi, \tag{17.30}$$

т. е. параметрическое представление некоторой кривой Γ. В этом смысле можно говорить, что уравнение (17.29) задает в полярных координатах кривую Γ. Для вычисления кривизны, радиуса кривизны и эволюты кривой Γ, заданной уравнением (17.29), надо перейти к ее параметрическому представлению (17.30) и воспользоваться выведенными выше формулами.

УПРАЖНЕНИЯ. 2. Пусть в полярных координатах задана кривая $\rho = \rho(\varphi)$, пусть α — угол наклона ее касательной к оси Ox, а ω — угол, образованный этой касательной с продолжением радиуса-вектора точки касания. Доказать, что $\alpha = \omega + \varphi$ и $\operatorname{tg} \omega = \frac{\rho}{\rho'}$.

3. Найти эволюту кривой $\rho = a(1 + \cos \varphi)$, $0 \leqslant \varphi \leqslant 2\pi$, называемой кардиоидой.

Указание. Воспользоваться результатами упражнений 1 и 2.

Задача 14. Пусть Γ — дважды дифференцируемая кривая без особых точек, $\Gamma = \{r(t);\ a \leqslant t \leqslant b\}$, и пусть $t_0 \in [a, b]$, $t_0 + \Delta t_1 \in [a, b]$, $t_0 + \Delta t_2 \in \in [a, b]$. Проведем через точки $r(t_0)$, $r(t_0 + \Delta t_1)$ и $r(t_0 + \Delta t_2)$ плоскость; доказать, что если в точке $r(t_0)$ кривизна $k \neq 0$, то при $\Delta t_1 \to 0$ и $\Delta t_2 \to 0$ эта плоскость стремится (определите это понятие) к соприкасающейся плоскости в точке $r(t_0)$.

Задача 15. В предположении предыдущей задачи проведем через те же три точки $r(t_0)$, $r(t_0 + \Delta t_1)$ и $r(t_0 + \Delta t_2)$ окружность. Доказать, что эта окружность при $\Delta t_1 \to 0$ и $\Delta t_2 \to 0$ стремится к окружности (опреде-

лите это понятие), лежащей в соприкасающейся плоскости с центром в центре кривизны кривой и радиусом, равным радиусу кривизны в точке $r(t_0)$.

Эта предельная окружность называется *соприкасающейся окружностью* в данной точке кривой.

17.6. Эвольвента

Как известно, $\frac{dt}{ds} = k\boldsymbol{n}$. Покажем, что для плоских кривых

$$\frac{d\boldsymbol{n}}{ds} = -k\boldsymbol{t}. \qquad (17.31)$$

В самом деле, поскольку $\boldsymbol{n}$ — единичный вектор и, следовательно, имеет постоянную длину, его производная $\frac{d\boldsymbol{n}}{ds}$ перпендикулярна ему. Касательный вектор $\boldsymbol{t}$ также перпендикулярен вектору $\boldsymbol{n}$. На плоскости два вектора, перпендикулярные третьему, коллинеарны, поэтому

$$\frac{d\boldsymbol{n}}{ds} = \alpha\boldsymbol{t}. \qquad (17.32)$$

Для того чтобы найти значение коэффициента α, продифференцируем по длине дуги тождество $\boldsymbol{tn} = 0$. В результате получим

$$\frac{d\boldsymbol{t}}{ds}\boldsymbol{n} + \boldsymbol{t}\frac{d\boldsymbol{n}}{ds} = 0.$$

Подставив сюда значения $\frac{d\boldsymbol{t}}{ds} = k\boldsymbol{n}$, $\frac{d\boldsymbol{n}}{ds} = \alpha\boldsymbol{t}$ и заметив, что $\boldsymbol{tt} = \boldsymbol{nn} = 1$, получим $\alpha = -k$. Отсюда, в силу равенства (17.32), и следует формула (17.31).

Формулы (17.9) и (17.31), т. е.

$$\frac{d\boldsymbol{t}}{ds} = k\boldsymbol{n}, \quad \frac{d\boldsymbol{n}}{ds} = -k\boldsymbol{t},$$

называются формулами Френе[1] для плоской кривой.

Определение 8. *Если кривая Γ_1 является эволютой плоской кривой Γ, то кривая Γ называется эвольвентой кривой Γ_1.*

---

[1] Ж. Ф. Френе (1816—1900) — французский математик.

При изучении взаимных свойств эволют и эвольвент ограничимся случаем, когда рассматриваемая плоская кривая Γ трижды непрерывно дифференцируема, задана своим представлением $\boldsymbol{r} = \boldsymbol{r}(s)$, $0 \leqslant s \leqslant S$, где s — длина ее дуги, радиус кривизны R кривой Γ не обращается ни в нуль, ни в бесконечность и всюду на кривой выполняется неравенство $\frac{dR}{ds} \neq 0$.

1^0. *Нормаль к эвольвенте является касательной к эволюте.*

Доказательство. Уравнение эволюты кривой имеет вид (см. п. 17.4)

$$\rho = \boldsymbol{r} + R\,\boldsymbol{n}. \tag{17.33}$$

Из этой формулы следует, что если кривая Γ является трижды непрерывно дифференцируемой кривой, то ее эволюта Γ_1 — непрерывно дифференцируемая кривая. Примем длину дуги s кривой Γ за параметр на эволюте Γ_1 и продифференцируем по нему уравнение эволюты (17.33):

$$\frac{d\rho}{ds} = \frac{d\boldsymbol{r}}{ds} + \frac{dR}{ds}\boldsymbol{n} + R\frac{d\boldsymbol{n}}{ds}. \tag{17.34}$$

Заметив, что

$$\frac{d\boldsymbol{r}}{ds} = \boldsymbol{t},\ \ R\frac{d\boldsymbol{n}}{ds} \underset{(17.31)}{=} -Rk\boldsymbol{t} = -\boldsymbol{t}, \tag{17.35}$$

так как $Rk = 1$; подставив выражения (17.35) для $\frac{d\boldsymbol{r}}{ds}$ и $R\frac{d\boldsymbol{n}}{ds}$ в равенство (17.34), получим

$$\frac{d\rho}{ds} = \frac{dR}{ds}\boldsymbol{n}. \tag{17.36}$$

Таким образом, вектор $\frac{d\rho}{ds}$, касательный к эволюте Γ_1, коллинеарен с вектором $\boldsymbol{n}$, нормальным к эвольвенте. Поэтому прямая, касательная к эволюте в центре кривизны некоторой точки эвольвенты, совпадает с нормальной прямой, проходящей через эту точку, так как обе эти прямые проходят через указанный центр кривизны (см. рис. 90, 91).

2^0. *Приращение длины дуги эволюты равно приращению радиуса кривизны эвольвенты.*

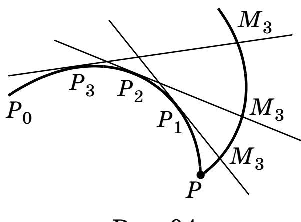


Рис. 94

Доказательство. Из равенства (17.36) следует, что

$$\left|\frac{d\rho}{ds}\right| = \left|\frac{dR}{ds}\right|. \qquad (17.37)$$

Обозначим через σ длину дуги эволюты Γ_1 кривой Γ, отсчитываемой в направлении возрастания дуги s самой кривой Γ. Тогда

$$\left|\frac{d\rho}{ds}\right|_{17.23} = \frac{d\sigma}{ds}. \qquad (17.38)$$

Если для определенности положить, что на рассматриваемой части кривой Γ ее радиус кривизны возрастает, то на этой части будет выполняться неравенство $\frac{dR}{ds} \geqslant 0$; тогда из равенств (17.37) и (17.38) следует, что

$$\frac{d\sigma}{ds} = \frac{dR}{ds}. \qquad (17.39)$$

Из равенства производных двух функций вытекает, что эти функции отличаются на константу:

$$\sigma(s) = R(s) + c, \qquad (17.40)$$

где c — некоторая постоянная. Приращение постоянной равно нулю, поэтому отсюда сразу имеем, что приращение длины дуги эволюты $\Delta\sigma = \sigma(s + \Delta s) - \sigma(s)$ совпадает с соответствующим приращением радиуса кривизны $\Delta R = R(s + \Delta s) - R(s)$, т. е. $\Delta\sigma = \Delta R$. □

Свойства 1^0 и 2^0 эволюты и эвольвенты имеют изящную механическую интерпретацию. Представим себе, что на кривую Γ_1 от точки P_0 до P натянута гибкая нерастяжимая нить, закрепленная в точке P_0. Если, туго натягивая эту нить, сматывать ее с кривой Γ_1, то конец P опишет кривую Γ, являющуюся эвольвентой кривой Γ_1 (рис. 94), так как длина дуги P_1P_2 кривой Γ_1 равна приращению длины отрезка P_1M_1 прямой, касательной к кривой Γ_1 в точке P_1:

$$\left|\overset{\frown}{P_1P_2}\right| = P_2M_2 - P_1M_1.$$

Таким образом, эвольвента кривой Γ_1 получается как бы развертыванием этой кривой, поэтому эвольвенту кривой называют также ее разверткой. Из сказанного также следу-

ет, что эвольвента кривой описывается точкой прямой, катящейся без скольжения по этой кривой.

Свойство 2^0 эволюты и эвольвенты дает возможность вычислять длину дуг эволюты, если известны радиусы кривизн эвольвенты. Найдем этим методом длину одной арки циклоиды (см. примеры в п. 16.2, 16.4 и 16.5).

В примере 3 п. 17.5 было показано, что для радиуса кривизны R циклоиды $x = r(t - \sin t)$, $y = r(1 - \cos t)$, $0 \leqslant t \leqslant 2\pi$, справедлива формула

$$R = R(t) = 4r \sin \frac{t}{2}$$

и что эволютой циклоиды является та же самая циклоида, но несколько сдвинутая. Поэтому длина половины арки циклоиды, соответствующей изменению параметра от 0 до π (на ней радиус кривизны возрастает), равен $R(\pi) - R(0) = 4r$. Следовательно, длина всей арки циклоиды равна $8r$.

17.7. Кручение пространственной кривой

Плоские кривые полностью с точностью до положения в пространстве описываются своей кривизной. Именно в дифференциальной геометрии доказывается, что для всякой непрерывной неотрицательной функции $k(s)$, $0 \leqslant s \leqslant S$, можно построить единственную с точностью до ее положения в пространстве плоскую кривую, для которой заданная функция является кривизной (см.: *Рашевский П. К.* Курс дифференциальной геометрии. — М.: ГИТТЛ, 1956).

Пространственные же кривые полностью описываются с помощью кривизны и так называемого кручения. Для его определения введем понятие бинормали.

Рассмотрим пространственную кривую $\Gamma = \{\boldsymbol{r}(s);\ 0 \leqslant s \leqslant S\}$, где s — переменная длина дуги.

Определение 9. *Векторное произведение единичного касательного вектора $\boldsymbol{t}$ и главной нормали $\boldsymbol{n}$ в данной точке кривой называется бинормалью кривой в этой точке.*

Бинормаль обозначается через $\boldsymbol{b}$. Таким образом,

$$\boldsymbol{b} \stackrel{\text{def}}{=} \boldsymbol{t} \times \boldsymbol{n}. \tag{17.41}$$

Очевидно, что бинормаль определена в тех точках, в которых определена главная нормаль, т. е. в которых кривизна не равна нулю.

Тройка единичных взаимно перпендикулярных векторов $\boldsymbol{t}$, $\boldsymbol{n}$ и $\boldsymbol{b}$ называется основным репером или, менее точно, основным трехгранником кривой в данной точке.

Из свойств векторного произведения следует, что кроме соотношения (17.41) имеют место соотношения

$$\boldsymbol{t} = \boldsymbol{n} \times \boldsymbol{b}, \quad \boldsymbol{n} = \boldsymbol{b} \times \boldsymbol{t}. \qquad (17.42)$$

Пусть $\boldsymbol{r}(s)$ — трижды дифференцируемая векторная функция. Продифференцировав равенство (17.41), получим

$$\frac{d\boldsymbol{b}}{ds} = \frac{d\boldsymbol{t}}{ds} \times \boldsymbol{n} + \boldsymbol{t} \times \frac{d\boldsymbol{n}}{ds} \underset{(17.14)}{=} k\boldsymbol{n} \times \boldsymbol{n} + \boldsymbol{t} \times \frac{d\boldsymbol{n}}{ds} = \boldsymbol{t} \times \frac{d\boldsymbol{n}}{ds}.$$

Согласно определению векторного произведения, вектор $\boldsymbol{t} \times \frac{d\boldsymbol{n}}{ds}$ перпендикулярен вектору $\boldsymbol{t}$. Кроме того, вектор $\boldsymbol{b}$ — единичный, поэтому, согласно лемме 1 п. 17.1, вектор $\frac{d\boldsymbol{b}}{ds}$ ортогонален вектору $\boldsymbol{b}$. Итак, $\frac{d\boldsymbol{b}}{ds} \perp \boldsymbol{t}$, $\frac{d\boldsymbol{b}}{ds} \perp \boldsymbol{b}$.

Вектор главной нормали $\boldsymbol{n}$ также ортогонален векторам $\boldsymbol{t}$ и $\boldsymbol{b}$, следовательно, векторы $\frac{d\boldsymbol{b}}{ds}$ и $\boldsymbol{n}$ коллинеарны, т. е. существует такое число $\varkappa$, что

$$\frac{d\boldsymbol{b}}{ds} = -\varkappa\boldsymbol{n}. \qquad (17.43)$$

Определение 10. *Коэффициент $\varkappa$ в формуле* (17.43) *называется кручением кривой в данной точке.*

В отличие от кривизны кручение кривой может быть как положительным, так и отрицательным (и, конечно, нулем).

Из формулы (17.43) следует, что абсолютная величина кручения $\varkappa$ является угловой скоростью вращения бинормали, или, что то же самое, угловой скоростью вращения соприкасающейся плоскости (так как она перпендикулярна бинормали).

17.8. Формулы Френе

Найдем производную главной нормали:

$$\frac{d\boldsymbol{n}}{ds} \underset{(17.42)}{=} \frac{d\boldsymbol{b}}{ds} \times \boldsymbol{t} + \boldsymbol{b} \times \frac{d\boldsymbol{t}}{ds} \underset{(17.43),\,(17.8)}{=}$$

$$\underset{(17.43),\,(17.8)}{=} -\varkappa\boldsymbol{n} \times \boldsymbol{t} + \boldsymbol{b} \times k\boldsymbol{n} \underset{(17.41),\,(17.42)}{=} -k\boldsymbol{t} + \varkappa\boldsymbol{b}. \qquad (17.44)$$

Итак, для производных основного репера кривой имеют место формулы (см. (17.8), (17.43) и (17.44)) разложения их по векторам этого репера:

$$\frac{d\boldsymbol{t}}{ds} = k\boldsymbol{n}, \qquad (17.45)$$

$$\frac{d\boldsymbol{n}}{ds} = -k\boldsymbol{t} + \varkappa\boldsymbol{b}, \ \frac{d\boldsymbol{b}}{ds} = -\varkappa\boldsymbol{n}.$$

Таким образом, матрица этих разложений кососимметрична и имеет вид

$$\begin{pmatrix} 0 & k & 0 \\ -k & 0 & \varkappa \\ 0 & -\varkappa & 0 \end{pmatrix}.$$

Формулы (17.45) называются *формулами Френе*.

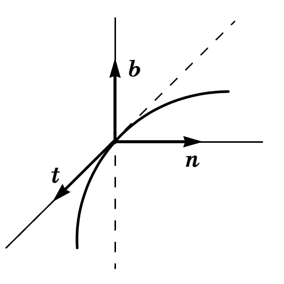


Рис. 95

Как известно, плоскость, проходящая через точку кривой параллельно касательному вектору и вектору главной нормали, т. е. перпендикулярно бинормали, называется соприкасающейся плоскостью.

Плоскость, проходящая через точку кривой параллельно главной нормали и бинормали, т. е. перпендикулярно касательному вектору, называется *нормальной плоскостью* кривой в этой точке.

Плоскость, проходящая через точку кривой параллельно касательной и бинормали, т. е. перпендикулярно главной нормали, называется *спрямляющей плоскостью* кривой в этой точке (рис. 95).

Для дальнейшего анализа геометрического смысла кривизны и кручения пространственной кривой в данной точке кривой поместим начало координат в эту точку, а векторы $\boldsymbol{t}$, $\boldsymbol{n}$ и $\boldsymbol{b}$ примем за единичные векторы координатных осей x, y, z. Если рассматриваемой точке отвечает значение длины дуги $s = s_0$, то разложим функцию $r(s)$ по формуле Тейлора в окрестности этой точки:

$$\boldsymbol{r}(s) = \boldsymbol{r}(s_0) + \frac{d\boldsymbol{r}(s_0)}{ds}\Delta s + \frac{1}{2}\frac{d^2\boldsymbol{r}(s_0)}{ds^2}\Delta s^2 + \frac{1}{6}\frac{d^3\boldsymbol{r}(s_0)}{ds^3}\Delta s^3 + \boldsymbol{o}(\Delta s^3),$$

где $\Delta s = s - s_0 \to 0$. Применив формулы Френе, получим

$$\Delta\boldsymbol{r} = \boldsymbol{r}(s) - \boldsymbol{r}(s_0) = \boldsymbol{t}\Delta s + \frac{1}{2}k\boldsymbol{n}\Delta s^2 + \frac{1}{6}\frac{d(n)}{ds}\Delta s^2 + \boldsymbol{o}(\Delta s^3) =$$

$$= \boldsymbol{t}\Delta s + \frac{1}{2}k\boldsymbol{n}\Delta s^2 + \frac{1}{6}\left(\frac{dk}{ds}\boldsymbol{n} - k^2\boldsymbol{t} + k\varkappa\boldsymbol{b}\right)\Delta s^3 + \boldsymbol{o}(\Delta s^3),\ \ \Delta s \to 0,$$

или в координатной записи (учитывая сделанный выбор координатной системы)

$$x = \Delta s - \frac{1}{6}k^2\Delta s^3 + \boldsymbol{o}(\Delta s^3), \tag{17.46}$$

$$y = \frac{1}{2}k\Delta s^2 + \frac{1}{6}\frac{dk}{ds}\Delta s^3 + \boldsymbol{o}(\Delta s^3), \tag{17.47}$$

$$z = \frac{1}{6}k\varkappa\Delta s^3 + \boldsymbol{o}(\Delta s^3). \tag{17.48}$$

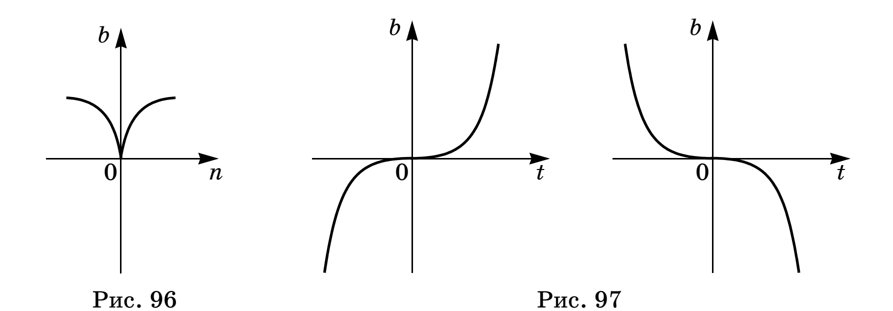


Рис. 96 Рис. 97

Рассмотрим проекцию кривой на ее соприкасающуюся плоскость, т. е. кривую, задаваемую уравнениями (17.46)—(17.47). Из (17.46) имеем $x \sim \Delta s$, $\Delta s \to 0$, поэтому $\Delta s = x + \boldsymbol{o}(\Delta s) = x + \boldsymbol{o}(x)$, $x \to 0$. Подставив выражение для Δs в уравнение (17.47), получим

$$y = \frac{1}{2}k(x + \boldsymbol{o}(x))^2 + \boldsymbol{o}(x^2) = \frac{1}{2}kx^2 + \boldsymbol{o}(x^2), \quad x \to 0.$$

Таким образом, проекция кривой на соприкасающуюся плоскость является с точностью до бесконечно малых высшего порядка параболой $y = \frac{1}{2}kx^2$. Кривизна кривой положительна, поэтому ветви этой параболы отходят от касательной в ту же сторону, в которую направлен вектор $\boldsymbol{n}$ (то, что вектор главной нормали указывает направление, в котором отклоняется кривая от своей касательной, было установлено и раньше, см. п. 17.3).

Рассмотрим теперь проекцию кривой на ее нормальную плоскость. Из (17.48) следует, что $z \sim \frac{1}{6}k\varkappa\Delta s^3$, $\Delta s \to 0$, откуда $\Delta s \sim \left(\frac{6z}{k\varkappa}\right)^{1/3}$, $z \to 0$, и, следовательно, $\Delta s = \left(\frac{6z}{k\varkappa}\right)^{1/3} + \boldsymbol{o}(z^{1/3})$, $z \to 0$. Подставив выражение для Δs в формулу (17.47), получим

$$y = \frac{1}{2}k\left(\left(\frac{6z}{k\varkappa}\right)^{2/3} + o(z^{2/3})\right) + o(z) = \left(\frac{9k}{2\varkappa^2}\right)^{1/3} z^{2/3} + o(z^{2/3}),\ z \to 0.$$

В этом случае проекция кривой с точностью до бесконечно малых высшего порядка является полукубической параболой (рис. 96) $y = \left(\frac{9k}{2\varkappa^2}\right)^{1/3} z^{2/3}$ и, следовательно, лежит по одну сторону от бинормали.

Для получения проекции на спрямляющую плоскость снова заметим, что из (17.46) следует, что $\Delta s = x + \boldsymbol{o}(x)$,

$x \to 0$. Подставив это выражение для Δs в равенство (17.48), имеем

$$z = \frac{1}{6}k\text{æ}x^3 + \boldsymbol{o}(x^3).$$

Таким образом, в этом случае проекция кривой с точностью до бесконечно малых более высокого порядка является кубической параболой

$$z = \frac{1}{6}k\text{æ}x^3,$$

поведение которой зависит от знака кручения (рис. 97). Если кручение $\text{æ} > 0$, то при возрастании параметра s кривая отходит от соприкасающейся плоскости в направлении бинормали $\boldsymbol{b}$, а при $\text{æ} < 0$ — в противоположном направлении (именно для того, чтобы имел место этот факт, в формуле (17.43) был взят знак минус, а не плюс).

17.9. Формулы для вычисления кручения

Уже были получены (см. п. 17.2) формулы

$$\frac{d\boldsymbol{r}}{ds} = \boldsymbol{t}, \quad \frac{d^2\boldsymbol{r}}{ds^2} = k\boldsymbol{n}. \tag{17.49}$$

Продифференцировав вторую из них, получим

$$\frac{d^3\boldsymbol{r}}{ds^3} = \frac{dk}{ds}\boldsymbol{n} + k\frac{d\boldsymbol{n}}{ds} \underset{(17.44)}{=} \frac{dk}{ds}\boldsymbol{n} - k^2\boldsymbol{t} + k\text{æ}\boldsymbol{b}. \tag{17.50}$$

Возьмем смешанное произведение векторов $\frac{d\boldsymbol{r}}{ds}$, $\frac{d^2\boldsymbol{r}}{ds^2}$ и $\frac{d^3\boldsymbol{r}}{ds^3}$

$$\left(\frac{d\boldsymbol{r}}{ds}, \frac{d^2\boldsymbol{r}}{ds^2}, \frac{d^3\boldsymbol{r}}{ds^3}\right) \underset{\substack{(17.49)\\(17.50)}}{=} \left(\boldsymbol{t}, k\boldsymbol{n}, \frac{dk}{ds}\boldsymbol{n} - k^2\boldsymbol{t} + k\text{æ}\boldsymbol{b}\right) = k^2\text{æ},$$

так как $\boldsymbol{n} \times \boldsymbol{n} = \boldsymbol{t} \times \boldsymbol{t} = \boldsymbol{0}$ и $(\boldsymbol{t}, \boldsymbol{n}, \boldsymbol{b}) = 1$.

Таким образом,

$$\text{æ} = \frac{\left(\frac{d\boldsymbol{r}}{ds}, \frac{d^2\boldsymbol{r}}{ds^2}, \frac{d^3\boldsymbol{r}}{ds^3}\right)}{k^2}, \tag{17.51}$$

и поскольку $k = \left|\frac{d^2\boldsymbol{r}}{ds^2}\right|$,

$$\text{æ} = \frac{\left(\frac{d\boldsymbol{r}}{ds}, \frac{d^2\boldsymbol{r}}{ds^2}, \frac{d^3\boldsymbol{r}}{ds^3}\right)}{\left|\frac{d^2\boldsymbol{r}}{ds^2}\right|}. \tag{17.52}$$

Для получения формул кручения при произвольном трижды дифференцируемом параметрическом задании $\boldsymbol{r}(t)$, $a \leqslant t \leqslant b$, кривой выразим производные $\frac{d\boldsymbol{r}}{ds}$, $\frac{d^2\boldsymbol{r}}{ds^2}$ и $\frac{d^3\boldsymbol{r}}{ds^3}$ через производные $\boldsymbol{r}'$, $\boldsymbol{r}''$ и $\boldsymbol{r}'''$ по параметру t:

$$\frac{d\boldsymbol{r}}{ds} = \boldsymbol{r}'\frac{dt}{ds}, \quad \frac{d^2\boldsymbol{r}}{ds^2} = \boldsymbol{r}''\left(\frac{dt}{ds}\right)^2 + \boldsymbol{r}'\frac{d^2t}{ds^2},$$

$$\frac{d^3\boldsymbol{r}}{ds^3} = \boldsymbol{r}'''\left(\frac{dt}{ds}\right)^3 + 3\boldsymbol{r}''\frac{dt}{ds}\frac{d^2t}{ds^2} + \boldsymbol{r}'\frac{d^3t}{ds^3}.$$

Подставив полученные значения производных по s формулу (17.51), имеем $\varkappa = \frac{(\boldsymbol{r}',\, \boldsymbol{r}'',\, \boldsymbol{r}''')}{k^2}\left(\frac{dt}{ds}\right)^6$, и так как $\left|\frac{ds}{dt}\right| = \left|\frac{d\boldsymbol{r}}{dt}\right| = |\boldsymbol{r}'|$, то

$$\varkappa = \frac{(\boldsymbol{r}',\, \boldsymbol{r}'',\, \boldsymbol{r}''')}{k^2|\boldsymbol{r}'|^6}. \qquad (17.53)$$

Заметив, что из формулы (17.10) для кривизны кривой следует, что $|\boldsymbol{r}'|^3 = \frac{|\boldsymbol{r}' \times \boldsymbol{r}''|}{k}$, получим из (17.53) еще одну формулу для кручения кривой:

$$\varkappa = \frac{(\boldsymbol{r}',\, \boldsymbol{r}'',\, \boldsymbol{r}''')}{|\boldsymbol{r}' \times \boldsymbol{r}''|^2}.$$

Можно показать, что кривизна и кручение полностью определяют форму пространственной кривой в том смысле, что для любых двух непрерывных функций $k(s) > 0$ и $\varkappa(s)$, $0 \leqslant s \leqslant S$, существует единственная кривая с точностью до ее положения в пространстве, для которой две данные функции являются соответственно кривизной и кручением как функциями длины дуги s кривой (см.: *Рашевский П. К.* Курс дифференциальной геометрии. — М.: ГИТТЛ, 1956).

УПРАЖНЕНИЯ. 4. Доказать, что если у кривой во всех точках ее кривизна равна нулю, то кривая является прямой.

5. Доказать, что если у кривой во всех точках ее кручение равно нулю, то кривая является плоской.

Глава 2

Интегральное исчисление функций одной переменной

§ 18.

Определения и свойства неопределенного интеграла

18.1. Первообразная и неопределенный интеграл

В этом параграфе рассматривается задача отыскания функции, для которой заданная функция является производной.

Пусть Δ — конечный или бесконечный промежуток числовой оси, т. е. интервал, полуинтервал или отрезок[1], и на Δ определены функции f и F.

Определение 1. *Функция F называется первообразной функцией (или, короче, первообразной) функции f на промежутке Δ, если F дифференцируема на Δ и в каждой точке этого промежутка производная функции F равна значению функции f:*

$$F'(x) = f(x),\ x \in \Delta. \tag{18.1}$$

При этом если некоторый конец промежутка Δ принадлежит промежутку, то под производной в этом конце понимается соответствующая односторонняя производная. Функция, имеющая в данной точке производную, непрерывна в этой точке (односторонне непрерывна, если речь идет об односторонней производной), поэтому первообразная F функции f непрерывна на промежутке Δ.

---

[1] Если рассматриваемый промежуток является отрезком, то само собой разумеется, что он может быть только конечным.

Пример. Функция $F(x) = \frac{x^3}{3}$ является первообразной функции $f(x) = x^2$ на всей числовой оси.

Иногда вместо «первообразная данной функции» говорят «первообразная для данной функции».

Первообразная любой функции, как это было отмечено, непрерывна. Функция же, у которой существует первообразная, не обязательно непрерывна, например, у разрывной функции

$$f(x) = \begin{cases} 2x \sin \frac{1}{x} - \cos \frac{1}{x} \text{ при } x \neq 0, \\ 0 \text{ при } x = 0, \end{cases}$$

на всей числовой оси существует первообразная

$$F(x) = \begin{cases} x^2 \sin \frac{1}{x} \text{ при } x \neq 0, \\ 0 \text{ при } x = 0. \end{cases}$$

(См. пример 9 в п. 9.7 и п. 14.2.)

ЛЕММА 1. *Две дифференцируемые на промежутке Δ функции F и Φ являются первообразными одной и той же функции в том и только том случае, когда они отличаются на постоянную*:

$$\Phi(x) = F(x) + C, \quad x \in \Delta, C = \text{const}. \tag{18.2}$$

Доказательство. Если F — первообразная функция f, т. е. $F' = f$, то и функция $F + C$ является первообразной той же функции f, так как $(F + C)' = F' = f$.

Если F и Φ — первообразные для одной и той же функции f, т. е. $F' = \Phi' = f$, то $(F - \Phi)' = F' - \Phi' = 0$ и, следовательно, согласно следствию 1 теоремы Лагранжа (см. теорему 3 п. 11.2), разность $F - \Phi = C$ — постоянная на промежутке Δ. □

Определение 2. *Пусть функция f определена на некотором промежутке. Совокупность всех ее первообразных на этом промежутке называется неопределенным интегралом от функции f и обозначается*

$$\int f(x)\, dx. \tag{18.3}$$

Символ $\int$ называется *знаком интеграла*, а $f(x)$ — подынтегральной функцией.

Если F — какая-либо первообразная функции f на рассматриваемом промежутке, то пишут

$$\int f(x)\,dx = F(x) + C, \qquad (18.4)$$

хотя правильнее было бы писать

$$\int f(x)\,dx = \{F(x) + C\}$$

(здесь и в дальнейшем C — произвольная постоянная).

Иногда под $\int f(x)\,dx$ понимается не совокупность всех первообразных функции f, а произвольный элемент этого множества, т. е. произвольная первообразная рассматриваемой функции.

Аналогичные ситуации уже встречались и раньше, например символом $f(x)$ обозначается как сама функция, так и ее значение в точке x. Из контекста обычно всегда бывает ясно, в каком смысле в данном месте употреблено то или иное обозначение.

Следует, однако, иметь в виду, *что всякое равенство, в обеих частях которого стоят неопределенные интегралы, есть равенство между множествами.*

Интеграл $\int f(x)\,dx$ есть совокупность первообразных функций f, поэтому, вместо того, чтобы сказать, что у функции f существует первообразная, говорят также, что существует интеграл $\int f(x)\,dx$.

Под знаком интеграла пишут для удобства не саму функцию f, а ее произведение на дифференциал dx. Это делается, например, для того чтобы указать, по какой переменной ищется первообразная:

$$\int x^2 z\,dx = \frac{x^3 z}{3} + C, \ \int x^2 z\,dz = \frac{x^2 z^2}{2} + C.$$

Здесь в обоих случаях подынтегральная функция равна $x^2 z$, но ее неопределенные интервалы в первом и втором случаях различны, так как в первом случае она рассматривается как функция от переменной x, а во втором — как функция от z.

Другие (более важные) соображения, показывающие целесообразность использования записи $\int f(x)\,dx$, будут ука-

заны в дальнейшем (см. замену переменного в интеграле, п. 18.3).

Если F — какая-либо первообразная функции f на промежутке Δ, то, согласно формуле (18.4), под знаком интеграла стоит дифференциал функции F:

$$dF(x) = F'(x)\,dx = f(x)\,dx.$$

По определению будем считать, что этот дифференциал под знаком интеграла можно записывать в любом из указанных видов, т. е. согласно этому соглашению,

$$\int f(x)\,dx = \int F'(x)\,dx = \int dF(x). \tag{18.5}$$

18.2. Основные свойства интеграла

Все рассматриваемые в этом пункте функции определены на некотором фиксированном промежутке Δ.

1^0. *Если функция F дифференцируема на некотором промежутке, то на нем $\int dF = F(x) + C$ или, что то же самое, $\int F'(x)\,dx = F(x) + C$.*

Это сразу следует из определения неопределенного интеграла как совокупности всех дифференцируемых функций, дифференциал которых стоит под знаком интеграла.

2^0. *Пусть функция f имеет первообразную на промежутке Δ, тогда для всех $x \in \Delta$ имеет место равенство*

$$d\int f(x)\,dx = f(x)\,dx. \tag{18.6}$$

Отметим, что в этом равенстве под интегралом $\int f(x)\,dx$ понимается произвольная первообразная F функции f. Поэтому равенство (18.6) можно записать в виде

$$dF(x) = f(x)\,dx,$$

справедливость последнего равенства следует из того, что F — первообразная f.

3^0. *Если функции f_1 и f_2 имеют первообразные на промежутке Δ, то и функция $f_1 + f_2$ имеет первообразную на этом промежутке, причем*

$$\int (f_1(x) + f_2(x))\,dx = \int f_1(x)\,dx + \int f_2(x)\,dx. \tag{18.7}$$

Это равенство выражает собой совпадение двух множеств функций и означает, что сумма каких-либо первообразных для функций f_1 и f_2 является первообразной для функции $f_1 + f_2$ и, наоборот, всякая первообразная для функции $f_1 + f_2$ является суммой некоторых первообразных для функций f_1 и f_2.

Свойство интеграла, выражаемое формулой (18.7), называется *аддитивностью интеграла относительно функций*.

Доказательство. Пусть F_1 и F_2 — первообразные соответственно функций f_1 и f_2, т. е. в каждой точке $x \in \Delta$ выполняются равенства $F_1'(x) = f_1(x)$, $F_2'(x) = f_2(x)$. Положим $F(x) = F_1(x) + F_2(x)$; тогда функция F является первообразной для функции $f_1 + f_2$, так как

$$F'(x) = F_1'(x) + F_2'(x) = f_1(x) + f_2(x), \quad x \in \Delta.$$

Следовательно, интеграл $\int (f_1(x) + f_2(x))\,dx$ состоит из функций $F(x) + C = F_1(x) + F_2(x) + C$, а сумма интегралов $\int f_1(x)\,dx + \int f_2(x)\,dx = F_1(x) + C_1 + F_2(x) + C_2$. Поскольку C, C_1 и C_2 — произвольные постоянные, оба эти множества, т. е. левая и правая части равенства (18.7), совпадают. □

4^0. *Если функция f имеет первообразную на промежутке Δ и k — число, то функция kf также имеет на Δ первообразную, причем при $k \neq 0$ справедливо равенство*

$$\int kf(x)\,dx = k\int f(x)\,dx. \tag{18.8}$$

Это равенство, также как равенство (18.8), является равенством множеств.

Доказательство. Пусть F — первообразная функции f, т. е. $F'(x) = f(x)$, $x \in \Delta$. Тогда функция kF является первообразной функции kf на промежутке Δ при любом $k \in \boldsymbol{R}$, так как $(kF(x))' = kF'(x) = kf(x)$, $x \in \Delta$. Поэтому интеграл $\int kf(x)\,dx$ состоит из всевозможных функций вида $kF + C$, а интеграл $k\int f(x)\,dx$ — из всевозможных функций $k(F + C) = kF + kC$. В силу произвольности постоянной C, при условии $k \neq 0$, обе совокупности функций совпадают. Это и означает справедливость равенства (18.8). □

СЛЕДСТВИЕ (линейность интеграла). *Если функции* f_1 *и* f_2 *имеют первообразные на промежутке* Δ, *а* λ_1 *и* λ_2 *— числа, то функция* $\lambda_1 f_1 + \lambda_2 f_2$ *также имеет первообразную на* Δ, *причем при* $\lambda_1^2 + \lambda_2^2 > 0$ *выполняется равенство*

$$\int (\lambda_1 f_1(x) + \lambda_2 f_2(x))\, dx = \lambda_1 \int f_1(x)\, dx + \lambda_2 \int f_2(x)\, dx.$$

Это непосредственно следует из свойств 3^0 и 4^0.

Вопрос о существовании первообразной будет рассмотрен ниже (см. п. 25.2), теперь же изучим простейшие методы вычисления первообразных для элементарных функций.

УПРАЖНЕНИЕ 1. Докажите, что для функции $f(x) = \operatorname{sign} x$ не существует такой функции F, что для всех $x \in \boldsymbol{R}$ выполняется равенство $F'(x) = \operatorname{sign} x$.

18.3. Табличные интегралы

Операция нахождения неопределенного интеграла от данной функции, называемая *интегрированием*, является действием, обратным дифференцированию, т. е. операции нахождения по данной функции ее производной (см. свойства 1^0 и 2^0 неопределенного интеграла в п. 18.2). Поэтому всякая формула, выражающая производную той или иной функции, т. е. формула вида $F'(x) = f(x)$, может быть обращена (записана в виде интегральной формулы):

$$\int f(x)\, dx = F(x) + C.$$

Используя это соображение, запишем таблицу значений ряда неопределенных интегралов, получающуюся непосредственно из соответствующей таблицы производных элементарных функций (см. § 9).

1. $\int x^\alpha\, dx = \dfrac{x^{\alpha+1}}{\alpha+1} + C$, $x > 0$, $\alpha \neq -1$.

Если число α таково, что степень x^α имеет смысл и для всех $x \leqslant 0$, то формула 1 справедлива на любом промежутке. Например, формула

$$\int x^2\, dx = \frac{x^3}{3} + C$$

справедлива на всей числовой оси.

Однако для интеграла $\int \frac{dx}{x^2}$ уже нельзя написать подобную единую формулу, справедливую для всей области определения подынтегральной функции, т. е. для всей числовой оси, из которой исключено число нуль. В этом случае имеем:

$$\int \frac{dx}{x^2} = \begin{cases} -\frac{1}{x} + C_1 \text{ для } x > 0, \\ -\frac{1}{x} + C_2 \text{ для } x < 0. \end{cases}$$

2. $\int \frac{dx}{x} = \ln |x| + C$

на любом промежутке, на котором $x \neq 0$.

3. $\int a^x \, dx = \frac{a^x}{\ln a} + C$, $a > 0$, $a \neq 1$.

В частности, $\int e^x \, dx = e^x + C$.

4. $\int \sin x \, dx = -\cos x + C$.

5. $\int \cos x \, dx = \sin x + C$.

6. $\int \frac{dx}{\cos^2 x} = \operatorname{tg} x + C$.

7. $\int \frac{dx}{\sin^2 x} = -\operatorname{ctg} x + C$.

8. $\int \operatorname{sh} x \, dx = \operatorname{ch} x + C$.

9. $\int \operatorname{ch} x \, dx = \operatorname{sh} x + C$.

10. $\int \frac{dx}{\operatorname{ch}^2 x} = \operatorname{th} x + C$.

11. $\int \frac{dx}{\operatorname{ch}^2 x} = -\operatorname{cth} x + C$.

12. $\int \frac{dx}{x^2 + a^2} = \frac{1}{a} \operatorname{arctg} \frac{x}{a} + C = -\frac{1}{a} \operatorname{arcctg} \frac{x}{a} + C$.

13. $\int \frac{dx}{x^2 - a^2} = \frac{1}{2a} \ln \left|\frac{x - a}{x + a}\right| + C$.

14. $\int \frac{dx}{\sqrt{a^2 - x^2}} = \arcsin \frac{x}{a} + C = -\arccos \frac{x}{a} + C$, $|x| < |a|$.

15. $\int \frac{dx}{\sqrt{x^2 \pm a^2}} = \ln \left|x + \sqrt{x^2 \pm a^2}\right| + C$,

причем, когда под корнем стоит $x^2 - a^2$, предполагается, что $|x| > |a|$.

Само собой разумеется, что если знаменатель подынтегральной функции обращается в нуль в некоторой точке, то написанные формулы будут справедливы лишь для тех промежутков, в которых не происходит обращения в нуль указанного знаменателя (см. формулы 2, 6, 7, 11, 13, 15). Это замечание относится и к аналогичным ситуациям, которые встретятся нам в дальнейшем и не будут каждый раз специально оговариваться.

То, что производными функций, стоящих в правых частях этих формул, являются соответствующие подынтегральные выражения, проверяется непосредственным дифференцированием (см. примеры в § 9).

С помощью интегралов 1—15, называемых обычно *табличными интегралами*, и доказанных выше свойств неопределенного интеграла можно выразить интегралы и от более сложных элементарных функций также через элементарные функции.

Например,

$$\int \left(5\cos x + 2 - 3x^2 + \frac{1}{x} - \frac{4}{x^2 + 1}\right) dx =$$

$$= 5\int \cos x\, dx + 2\int dx - 3\int x^2\, dx + \int \frac{dx}{x} - 4\int \frac{dx}{x^2 + 1} =$$

$$= 5 \sin x + 2x - x^3 + \ln |x| - 4 \operatorname{arctg} x + C.$$

Отметим, что для всякого многочлена степени n существует первообразная и она является многочленом степени $n + 1$, точнее,

$$\int (a_0 + a_1x + a_2x^2 + \dots + a_nx^n)\, dx =$$

$$= a_0x + \frac{a_1x^2}{2} + \frac{a_2x^3}{3} + \dots + \frac{a_nx^{n+1}}{n+1} + C. \qquad (18.9)$$

Это следует из свойств 3^0 и 4^0 неопределенного интеграла (см. п. 18.2) и формулы 1 этого пункта.

Если первообразная некоторой функции f является элементарной функцией, то говорят, что интеграл $\int f(x)\,dx$ выражается через элементарные функции или что этот интеграл вычисляется.

18.4. Интегрирование подстановкой (замена переменной)

В этом и следующем пунктах будут рассмотрены два свойства неопределенного интеграла, часто оказывающиеся полезными при вычислении первообразных элементарных функций.

ТЕОРЕМА 1. *Пусть функции $f(x)$ и $\varphi(t)$ определены соответственно на промежутках Δ_x и Δ_t, причем $\varphi(\Delta_t) \subset \Delta_x$. Если функция f имеет на Δ_x первообразную $F(x)$ и, следовательно,*

$$\int f(x)\,dx = F(x) + C, \tag{18.10}$$

а функция φ дифференцируема на Δ_t, то функция $f(\varphi(t))\varphi'(t)$ имеет на Δ_t первообразную $F(\varphi(t))$ и

$$\int f(\varphi(t))\varphi'(t)\,dt = \int f(x)\,dx\Big|_{x=\varphi(t)}. \tag{18.11}$$

Доказательство. Функции f и F определены на промежутке Δ_x, и так как, по условию теоремы, справедливо включение $\varphi(\Delta_t) \subset \Delta_x$, то имеют смысл сложные функции $f(\varphi(t))$ и $F(\varphi(t))$. При этом так как

$$F'(x) = f(x),\ x \in \Delta_x, \tag{18.12}$$

то по правилу дифференцирования сложной функции получим

$$\frac{d}{dt}F(\varphi(t)) = \frac{dF}{dx}\Big|_{x=\varphi(t)} \frac{d\varphi(t)}{dt} = f(\varphi(t))\varphi'(t), \quad t \in \Delta_t.$$

Это и означает, что функция $f(\varphi(t))\varphi'(t)$ имеет в качестве одной из своих первообразных функцию $F(\varphi(t))$. Отсюда, согласно определению интеграла, следует, что

$$\int f(\varphi(t))\varphi'(t)\,dt = F(\varphi(t)) + C. \tag{18.13}$$

Подставив же в формулу (18.10) $x = \varphi(t)$, получим

$$\int f(x)\,dx\Big|_{x=\varphi(t)} = F(\varphi(t)) + C. \tag{18.14}$$

В формулах (18.13) и (18.14) равны правые части, значит, равны и левые, т. е. имеет место равенство (18.11). □

Формула (18.11) называется формулой интегрирования подстановкой, а именно подстановкой $\varphi(t) = x$. Это название объясняется тем, что если формулу (18.11) записать в виде

$$\int f(\varphi(t))\, d\varphi(t) = \int f(x)\, dx\big|_{x = \varphi(t)}, \qquad (18.15)$$

то будет видно, что, для того чтобы вычислить интеграл $\int f(\varphi(t))\varphi'(t)\ dt = \int f(\varphi(t))\, d\varphi(t)$, можно сделать подстановку $x = \varphi(t)$, вычислить интеграл $\int f(x)\, dx$ и затем вернуться к переменной t, положив $x = \varphi(t)$.

Примеры. 1. Для вычисления интеграла $\int \cos ax\, dx$ естественно сделать подстановку $u = ax$, тогда

$$\int \cos ax\, dx = \frac{1}{a}\int \cos u\, du = \frac{1}{a}\sin u + C = \frac{1}{a}\sin ax + C,\ a \neq 0.$$

2. Для вычисления интеграла $\int \frac{x dx}{x^2 + a^2}$ удобно применить подстановку $u = x^3 + a^2$:

$$\int \frac{x dx}{x^2 + a^2} = \frac{1}{2}\int \frac{du}{u} = \frac{1}{2}\ln |u| + C = \frac{1}{2}\ln (x^2 + a^2) + C.$$

3. При вычислении интегралов вида $\int \frac{\varphi'(x)dx}{\varphi(x)}$, $\varphi(x) \neq 0$, полезна подстановка $u = \varphi(x)$:

$$\int \frac{\varphi'(x)}{\varphi(x)}\ dx = \int \frac{d\varphi(x)}{\varphi(x)} = \int \frac{du}{u} = \ln |\varphi(x)| + C.$$

Например,

$$\int \operatorname{tg} x\, dx = -\int \frac{d\cos x}{\cos x} = -\ln |\cos x| + C.$$

Иногда, прежде чем применить метод интегрирования подстановкой, приходится проделать более сложные преобразования подынтегральной функции:

$$\int \frac{dx}{\sin x} = \int \frac{dx}{2\sin\frac{x}{2}\cos\frac{x}{2}} = \int \frac{1}{\operatorname{tg}\frac{x}{2}}\frac{dx}{2\cos^2\frac{x}{2}} = \int \frac{d\operatorname{tg}\frac{x}{2}}{\operatorname{tg}\frac{x}{2}} = \ln \left|\operatorname{tg}\frac{x}{2}\right| + C.$$

Отметим, что формулу (18.11) бывает целесообразно использовать и в обратном порядке, т. е. справа налево. Именно, иногда удобно вычисление интеграла $\int f(x)\, dx$ с помощью

соответствующей замены переменного $x = \varphi(t)$ свести к вычислению интеграла $\int f(\varphi(t))\varphi'(t)\,dt$ (если этот интеграл в каком-то смысле «проще» исходного).

В случае, когда функция φ имеет обратную φ^{-1}, перейдя в обеих частях формулы (18.11) к переменной x с помощью подстановки $t = \varphi^{-1}(x)$ и поменяв местами стороны равенства, получим

$$\int f(x)\,dx = \int f(\varphi(t))\varphi'(t)\,dt\big|_{t = \varphi^{-1}(x)}.$$

Эта формула называется обычно *формулой интегрирования заменой переменной*.

Для того чтобы существовала функция φ^{-1}, обратная φ, в дополнение к условиям теоремы 1 достаточно, например, потребовать, чтобы на рассматриваемом промежутке Δ_t функция φ была строго монотонной. В этом случае, как известно (см. п. 6.3), существует однозначная обратная функция φ^{-1}.

Примеры. 4. Интегралы вида $\int \frac{dx}{\sqrt{ax^2 + bx + c}}$, $a \neq 0$, в том случае, когда подкоренное выражение неотрицательно на некотором промежутке[1], легко сводятся с помощью замены переменного к табличным.

Действительно, замечая, что $ax^2 + bx + c = a\left(x + \frac{b}{2a}\right)^2 + c - \frac{b^2}{4a}$, сделаем замену переменной $t = \sqrt{|a|}\left(x + \frac{b}{2a}\right)$ и положим $d = c - \frac{b^2}{4a}$. Тогда $dx = \frac{dt}{\sqrt{|a|}}$ и, в силу формулы (18.11), получим

$$\int \frac{dx}{\sqrt{ax^2 + bx + c}} = \frac{1}{\sqrt{|a|}} \int \frac{dt}{\sqrt{\pm t^2 + d}}$$

(перед t^2 стоит знак плюс, если $a > 0$, и знак минус, если $a < 0$). Интеграл, стоящий в правой части равенства, является табличным (см. формулы 14 и 15 в п. 18.3). Найдя его по

[1] В противном случае, т. е. когда подкоренное выражение отрицательно для всех $x \in \boldsymbol{R}$, получится интеграл от комплекснозначной функции. Такие интегралы здесь не рассматриваются.

соответствующим формулам и вернувшись от переменной t к переменной x, получим искомый интеграл.

Подобным же приемом вычисляются и интегралы вида

$$\int \frac{dx}{ax^2 + bx + c}, \quad a \neq 0$$

(см. об этом в п. 20.1).

5. Интеграл $\int \sqrt{a^2 - x^2}\, dx$ можно вычислить с помощью подстановки $x = a\sin t$ (см. также пример 2 в п. 18.5). Имеем $dx = a \cos t\, dt$, поэтому

$$\int \sqrt{a^2 - x^2}\, dx = a^2 \int \cos^2 t\, dt =$$

$$= a^2 \int \frac{1 + \cos 2t}{2}\, dt = \frac{a^2}{2} \int dt + \frac{a^2}{2} \int \cos 2t\, dt = \frac{a^2 t}{2} + \frac{a^2}{4} \sin 2t + C.$$

Подставляя это выражение $t = \arcsin \frac{x}{a}$ и замечая, что

$$\sin 2 \arcsin \frac{x}{a} = 2 \sin \left(\arcsin \frac{x}{a}\right) \cos \left(\arcsin \frac{x}{a}\right) =$$

$$= 2\frac{x}{a}\sqrt{1 - \frac{x^2}{a^2}} = \frac{2}{a^2} x \sqrt{a^2 - x^2},$$

окончательно будем иметь

$$\int \sqrt{a^2 - x^2}\, dx = \frac{a^2}{2} \arcsin \frac{x}{a} + \frac{x}{2}\sqrt{a^2 - x^2} + C.$$

Заметим, что для проверки результата, полученного при вычислении неопределенного интеграла, достаточно его продифференцировать, после чего должно получиться подынтегральное выражение вычисляемого интеграла.

Другие примеры на интегрирование с помощью замены переменного рассмотрены в § 21, 22.

18.5. Интегрирование по частям

ТЕОРЕМА 2. *Если функции $u(x)$ и $v(x)$ дифференцируемы на некотором промежутке и на этом промежутке существует интеграл $\int v\,du$, то на нем существует и интеграл $\int u\,dv$, причем*

$$\int u\,dv = uv - \int v\,du. \tag{18.16}$$

Доказательство. Пусть функции u и v дифференцируемы на промежутке Δ; тогда по правилу дифференцирования произведения для всех точек этого промежутка имеет место равенство

$$d(uv) = vdu + udv,$$

поэтому

$$udv = d(uv) - vdu.$$

Интеграл от каждого слагаемого правой части существует, так как, согласно свойству 1^0 п. 18.2,

$$\int d(uv) = uv + C,$$

а интеграл $\int vdu$ существует по условию теоремы. Поэтому на основании свойства 3^0 п. 18.2 существует и интеграл $\int udv$, причем

$$\int udv = \int d(uv) - \int vdu. \tag{18.17}$$

Подставляя в правую часть (18.17) $uv + C$ вместо $\int d(uv)$ и относя произвольную постоянную C к интегралу $\int vdu$, получим формулу (18.16). □

С помощью формулы (18.16) вычисляются многие интегралы. При ее практическом использовании задана левая часть (18.16), т. е. функция u и дифференциал dv, а поэтому v определяется неоднозначно. Обычно в качестве v выбирается функция, записываемая наиболее простой формулой.

Примеры. 1. Пусть требуется вычислить интеграл $\int xe^x\,dx$. Полагая $u = x$, $dv = e^x\,dx$, откуда $du = dx$, $v = e^x$, имеем

$$\int xe^x\,dx = \int x\,de^x = xe^x - \int e^x\,dx = xe^x - e^x + C.$$

Заметим, что, взяв $u = e^x$ и $dv = xdx$, откуда $du = e^x\,dx$ и $v = x^2/2$, мы имели бы

$$\int xe^x\,dx = \frac{1}{2}x^2e^x - \frac{1}{2}\int x^2e^x\,dx,$$

т. е. интегрирование по частям привело бы к интегралу, более сложному, чем исходный. Отсюда видно, что при вычислении интегралов с помощью формулы (18.14) не каждый

способ выбора функций u и v приводит к интегралу, более простому, чем первоначальный.

2. Вычислим интеграл $I = \int \sqrt{a^2 - x^2}\ dx$ посредством интегрирования по частям (ранее, см. п. 18.4, пример 5, он был вычислен с помощью замены переменного).

Полагая $u = \sqrt{a^2 - x^2}$, $dv = dx$ и, следовательно, $du = -\frac{x}{\sqrt{a^2 - x^2}}\ dx$, $v = x$, получим

$$I = \int \sqrt{a^2 - x^2}dx = x\sqrt{a^2 - x^2} + \int \frac{x^2dx}{\sqrt{a^2 - x^2}}. \qquad (18.18)$$

Добавим и вычтем a^2 в числителе подынтегральной функции интеграла, стоящего в правой части равенства; тогда, произведя деление на $\sqrt{a^2 - x^2}$, будем иметь

$$\int \frac{x^2dx}{\sqrt{a^2 - x^2}} = \int \frac{a^2 - (a^2 - x^2)}{\sqrt{a^2 - x^2}}\ dx =$$

$$= a^2\int \frac{dx}{\sqrt{a - x^2}} - \int \sqrt{a^2 - x^2}\ dx = a^2\arcsin \frac{x}{a} - I.$$

Подставив это выражение в (18.8), получим

$$I = x\sqrt{a^2 - x^2} + a^2\arcsin \frac{x}{a} - I. \qquad (18.19)$$

Как уже отмечалось, всякое равенство такого вида представляет собой равенство между двумя множествами функций, элементы каждого из которых отличаются друг от друга на постоянную. Поэтому общее выражение для элемента множества I, согласно (18.19), имеет вид

$$I = \frac{x}{2}\sqrt{a^2 - x^2} + \frac{a^2}{2}\arcsin \frac{x}{a} + C.$$

3. Иногда для вычисления интеграла правило интегрирования по частям приходится применять несколько раз, например,

$$\int \arcsin^2 x\ dx = x\arcsin^2 x - 2\int \arcsin x \frac{xdx}{\sqrt{1 - x^2}} =$$

$$= x\arcsin^2 x + 2\int \arcsin x\ d\sqrt{1 - x^2} =$$

$$= x\arcsin^2 x + 2\arcsin x\sqrt{1 - x^2} - 2x + C.$$

4. Если $P_n(x)$ — многочлен степени n, то для вычисления интеграла $\int P_n(x)e^{\alpha x}\,dx$ следует формулу интегрирования по частям применить n раз. Выполнив это, получим

$$\int P_n(x)e^{\alpha x}\,dx = e^{\alpha x}\left(\frac{P_n(x)}{\alpha} - \frac{P_n'(x)}{\alpha^2} + \ldots + (-1)^n \frac{P_n^{(n)}(x)}{\alpha^{n+1}}\right) + C.$$

Другие примеры на применение интегрирования по частям будут рассмотрены в § 22.

18.6*. Обобщение понятия первообразной

Введенные выше понятия первообразной и интеграла могут быть обобщены в разных направлениях, причем таким образом, что будут сохраняться как свойства 1^0—4^0 п. 18.2, так и теоремы 1 и 2. Ограничимся здесь лишь одним обобщением, которое нам понадобится в дальнейшем, например при интегрировании так называемых кусочно-непрерывных функций.

Пусть функции f и F определены на некотором конечном или бесконечном промежутке Δ.

Определение 3. *Функция F называется обобщенной первообразной функции f на Δ, если*:

1) *функция F непрерывна на промежутке* Δ;

2) *в любой точке промежутка Δ, кроме некоторого конечного множества $E_f = E_{f,\,F} \subset \Delta$, функция F имеет производную, равную значению функции f в этой точке*:

$$F'(x) = f(x),\ x \in \Delta E_f. \qquad (18.20)$$

В точках множества E_f функция F, будучи согласно условию 1 обязательно непрерывной, может иметь или не иметь производной, причем, если производная существует, то $F'(x) \neq f(x)$.

В частности, множество E_f может быть пустым. В этом случае понятие обобщенной первообразной совпадает с понятием первообразной в смысле определения 1.

В этом пункте рассматривается только обобщенная первообразная в смысле определения 3, поэтому для краткости она будет, как правило, называться первообразной, т. е. каждый раз эпитет «обобщенная» не будет употребляться.

Пример. Для функции $f(x) = \text{sign}\, x$ функция $F(x) = |x|$, $-\infty < x < +\infty$, является первообразной на всей числовой оси. Здесь множество E_f состоит из одной точки — нуля: $E_{\text{sign}\, x} = \{0\}$. Для любой же точки $x \neq 0$ имеет место равенство $|x|' = \text{sign}\, x$.

Отметим, что функция $F(x) = |x|'$ является первообразной и для функции

$$f_1(x) = \begin{cases} 1 & \text{при } x \geqslant 0, \\ -1 & \text{при } x < 0, \end{cases}$$

отличающейся от функции $f(x) = \text{sign}\, x$ значением в нуле.

На этом примере видно, что одна и та же функция F может быть первообразной для разных функций f, однако, в силу условия (18.20), эти функции f могут отличаться друг от друга только значениями на конечном множестве точек (зависящем от выбираемых функций f).

Очевидно, что если функция F является первообразной функции f на некотором промежутке Δ, т. е. функция F непрерывна на Δ и во всех его точках, кроме некоторого конечного множества, выполняется условие $F'(x) = f(x)$, то для любой постоянной C функция $F(x) + C$ также непрерывна на промежутке Δ и во всех его точках, кроме указанного выше конечного их множества, выполняется условие

$$(F(x) + 4\, C)' = F'(x) + C' = f(x),$$

т. е. функция $F(x) + C$ также является первообразной функции на промежутке Δ.

С другой стороны, если функции F и Φ являются первообразными для функции f на промежутке Δ, т. е. если F и Φ непрерывны на Δ и во всех его точках, кроме конечных множеств $E_{f,\, F}$ и соответственно $E_{f,\, \Phi}$, выполняются условия

$$F'(x) = f(x) \text{ и } \Phi'(x) = f(x),$$

то для всех точек промежутка Δ, кроме множества $E_{f,\, F} \cup E_{f,\, \Phi}$, будет выполняться условие

$$F'(x) = f(x) = \Phi'(x),$$

причем множество $E_{f,\, F} \cup E_{f,\, \Phi}$, где это условие нарушается, конечное как объединение двух конечных множеств.

Отсюда, в силу следствия 1 теоремы 3 из п. 11.2, вытекает, что функции F и Φ отличаются на промежутке Δ лишь на некоторую постоянную C:

$$\Phi(x) = F(x) + C,\ x \in \Delta. \tag{18.21}$$

Таким образом, совокупность обобщенных первообразных одной и той же функции, как и совокупность обычных первообразных в смысле определения 1, состоит из функций, отличающихся друг от друга на постоянную.

По аналогии с определением 2 вводится в рассматриваемом случае и понятие интеграла.

Определение 4. *Совокупность всех обобщенных первообразных функции f, заданной на некотором промежутке Δ, называется неопределенным интегралом от функции f на этом промежутке и обозначается тем же символом, что и раньше*:

$$\int f(x)\,dx.$$

Если F — первообразная функции f на промежутке Δ, то, согласно определению 4, в формуле (18.3) под знаком интеграла стоит дифференциал функции F в точках $x \in \Delta \backslash E_{f,\,F}$:

$$\int f(x)\,dx = \int F'(x)\,dx = \int dF\,(x). \tag{18.22}$$

Покажем, что для интеграла, состоящего из обобщенных первообразных, сохраняются свойства интеграла, доказанные в п. 18.2, 18.4 и 18.5.

1^0_1. *Пусть функция F непрерывна на промежутке Δ и дифференцируема во всех его точках, кроме некоторого конечного их множества $E \subset \Delta$, тогда*

$$\int dF(x) = F(x) + C,$$

или, что то же самое (см. 18.20)),

$$\int F'(x)\,dx = F(x) + C.$$

Справедливость этого равенства вытекает из определения неопределенного интеграла как совокупности всех функций, непрерывных на данном промежутке Δ, дифференциал которых во всех точках $x \in \Delta$, кроме конечного их множества (своего для каждой функции F) стоит под знаком интеграла (см. 18.20)), и из общего вида (18.21) всех первообразных данной функции.

2^0_1. *Пусть функция f имеет первообразную на промежутке Δ. Тогда для любой точки $x \in \Delta \backslash E_f$ имеет место равенство*

$$d\int f(x)\,dx = f(x)\,dx. \tag{18.23}$$

Как и в п. 18.1, здесь под интегралом $\int f(x)\,dx$ понимается произвольная первообразная функции f. Справедливость формулы (18.23) очевидна в силу определения первообразной.

3_1^0. *Если функции f_1 и f_2 имеют первообразные на некотором промежутке, то и функция $f_1 + f_2$ также имеет первообразную на этом промежутке, причем*

$$\int (f_1(x) + f_2(x))\,dx = \int f_1(x)\,dx + \int f_2(x)\,dx. \qquad (18.24)$$

Доказательство. Пусть функции f_1 и f_2 имеют на некотором промежутке Δ первообразные F_1 и F_2, следовательно, $\int f_1(x)\,dx = F_1(x) + C_1$, $\int f_2(x)\,dx = F_2(x) + C_2$. Функции F_1и F_2 непрерывны на промежутке Δ, а в его точках $x \in \Delta\backslash E_{f_1}$ и $x \in \Delta\backslash E_{f_2}$ соответственно выполняются условия $F_1'(x) = f_1(x)$ и $F_2'(x) = f_2(x)$, где E_{f_1} и E_{f_2} — некоторые конечные множества.

Положим $F = F_1 + F_2$. Тогда функция F непрерывна на промежутке Δ как сумма непрерывных на этом промежутке функций F_1 и F_2 и для любой точки $x \notin \Delta\backslash(E_{f_1} \cup E_{f_2})$, $x \in \Delta$ имеет место равенство

$$F'(x) = [F_1(x) + F_2(x)]' = F_1'(x) + F_2'(x) = f_1(x) + f_2(x),$$

причем множество $E_{f_1} \cup E_{f_2}$, для точек которого это равенство нарушается, является конечным, как объединение двух конечных множеств E_{f_1} и E_{f_2}. Это означает, что F является первообразной для функции $f_1 + f_2$ на Δ, следовательно,

$$\int (f_1(x) + f_2(x))\,dx = F(x) + C = F_1(x) + F_2(x) + C.$$

Таким образом, левая часть формулы (18.24) состоит из функций вида $F_1(x) + F_2(x) + C$, правая — из функций вида $F_1(x) + C_1 + F_2(x) + C_2$. Так как постоянные C, C_1 и C_2 произвольные, то эти совокупности совпадают. □

4_1^0. *Если функция f имеет на некотором промежутке первообразную и k — число, то функция kf также имеет*

на этом промежутке первообразную, причем при $k \neq 0$ *справедливо равенство*

$$\int kf(x)\,dx = k\int f(x)\,dx. \qquad (18.25)$$

Доказательство. Пусть на некотором промежутке Δ имеем $\int f(x)\,dx = F(x) + C$, т. е. функция F непрерывна на Δ и во всех точках $x \in \Delta$, кроме конечного множества E_f, выполняется условие

$$F'(x) = f(x).$$

Тогда функция kF также непрерывна на Δ и во всех точках $x \in \Delta \setminus E_f$ имеет место равенство

$$(kF(x))' = kF'(x) = kf(x).$$

Это означает, что функция kF является первообразной для kF, следовательно,

$$\int kf(x)\,dx = kF(x) + C_1.$$

Таким образом, левая часть формулы (18.25) представляет собой совокупность функций вида $kF(x) + C_1$, а правая состоит из функций вида $k(F(x) + C) = kF(x) + kC$. Постоянные C и C_1 произвольны, поэтому при условии $k \neq 0$ эти совокупности совпадают. □

ТЕОРЕМА 3. *Пусть функции* $f(x)$ *и* $\varphi(x)$ *определены соответственно на промежутках* Δ_x *и* Δ_t, $\varphi(\Delta_t) \subset \Delta_x$, *функция* f *имеет на* Δ_x *первообразную* $F(x)$ *и, следовательно,*

$$\int f(x)\,dx = F(x) + C, \qquad (18.26)$$

E_f *— такое конечное множество, что* $E_f \subset \Delta_x$, *и для всех* $x \in \Delta_x \setminus E_f$ *выполняется равенство* $F'(x) = f(x)$.

Если функция φ *непрерывна на промежутке* Δ_t, *дифференцируема во всех его точках, за исключением некоторого конечного множества, и полный прообраз* $\varphi^{-1}(E_f)$ *множества* E_f *также является конечным множеством, то функция* $f(\varphi(t))\varphi'(t)$ *имеет на промежутке* Δ_t *первообразную* $F(\varphi(t))$, *и поэтому*

$$\int f(\varphi(t))\varphi'(t)\,dt = F(\varphi(t)) + C \underset{(18.26)}{=} \int f(x)\,dx\Big|_{x = \varphi(t)}. \qquad (18.27)$$

Доказательство. Функции $f(x)$ и $F(x)$ определены на промежутке Δ_x и по условию теоремы справедливо включение $\varphi(\Delta_t) \subset \Delta_x$, поэтому имеют смысл сложные функции $f(\varphi(t))$ и $F(\varphi(t))$. Согласно условиям теоремы, функция φ непрерывна на промежутке Δ_t и существует такое конечное множество (обозначим его через E_φ), что функция φ дифференцируема во всех точках $t \in \Delta_t \backslash E_\varphi$. Следовательно, функция $F(\varphi(t))$ непрерывна на Δ_t как композиция непрерывных функций и, по правилу дифференцирования сложных функций, для всех точек $t \in (\Delta_t \backslash E_\varphi \cup \varphi^{-1}(E_f))$ имеет место равенство

$$\frac{d}{dt}F(\varphi(t)) = \left.\frac{dF(x)}{dx}\right|_{x = \varphi(t)} \frac{d\varphi(t)}{dt} = f(\varphi(t))\varphi'(t),$$

причем множество $E_\varphi \cup \varphi^{-1}(E_f)$, где указанное равенство может не иметь места, является конечным множеством как сумма двух конечных множеств: E_φ и $\varphi^{-1}(E_f)$. Это и означает, что функция $f(\varphi(t))\varphi'(t)$ имеет в качестве одной из своих первообразных функцию $F(\varphi(t))$. Отсюда сразу и следует формула (18.27). □

ТЕОРЕМА 4. *Если каждая из функций $u(x)$ и $v(x)$, заданных на данном промежутке, дифференцируема во всех его точках, кроме конечного их множества, и на этом промежутке существует интеграл $\int vdu$, то на том же промежутке существует и интеграл $\int udv$, причем*

$$\int udv = uv - \int vdu. \qquad (18.28)$$

Доказательство. Пусть функции $u(x)$ и $v(x)$ заданы на промежутке Δ, причем $u(x)$ не дифференцируема только на конечном множестве $E_u \subset \Delta$, а $v(x)$ не дифференцируема на конечном множестве $E_v \subset \Delta$ и $E = E_u \cup E_v$. Очевидно, что E — также конечное множество и что для всех точек $x \in \Delta \backslash E$, по правилу дифференцирования произведения, будем иметь

$$d(uv) = vdu + udv,$$

поэтому

$$udv = d(uv) - vdu.$$

Интеграл от каждого слагаемого правой части этого равенства существует, так как, согласно свойству 1_1^0,

$$\int d\,(uv) = uv + C,$$

а интеграл $\int v\,du$ существует по условию теоремы. Поэтому на основании свойства 3_1^0 существует и интеграл $\int udv$, причем

$$\int udv = \int d\,(uv) - \int vdu.$$

Подставляя в правую часть этого равенства $uv + C$ вместо $\int d(uv)$ и относя произвольную постоянную C к интегралу $\int vdu$, получим формулу (18.28). □

§ 19.

Некоторые сведения о комплексных числах и многочленах

19.1. Комплексные числа

Как известно из алгебры, *комплексными числами* называются выражения вида

$$z = x + yi,$$

где i — некоторый элемент, называемый *мнимой единицей*, а x и y — любые действительные числа. Множество всех комплексных чисел обозначается через $\boldsymbol{C}$. Число x называется *действительной частью*, y — *мнимой частью* комплексного числа $z = x + yi$. Это записывается следующим образом: $x = \operatorname{Re} z$, $y = \operatorname{Im} z$[1].

Обозначение $x + iy$ комплексных чисел называется их *алгебраической формой* записи.

Комплексное число z, не являющееся действительным, т. е. у которого $\operatorname{Im} z \neq 0$, будем называть *существенно комплексным числом*. Число $\sqrt{x^2 + y^2}$ называется *модулем* комплексного числа $z = x + iy$ и обозначается $|z|$, т. е. $|z| = \sqrt{x^2 + y^2}$.

[1] От лат. realis — действительный и imaginarius — мнимый.

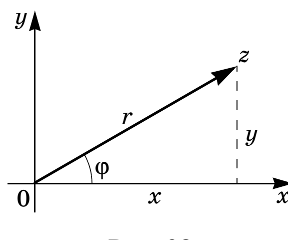


Рис. 98

Каждому комплексному числу $z = x + iy$ соответствует упорядоченная пара действительных чисел (x, y), и обратно; каждой упорядоченной паре действительных чисел (x, y) соответствует комплексное число $z = x + yi$. В силу этого взаимно однозначного соответствия (а также и в силу других обстоятельств, о которых речь будет и идти ниже) комплексное число $z = x + yi$ геометрически удобно интерпретировать либо как точку (x, y), либо как радиус-вектор на плоскости с координатами x и y (при некоторой фиксированной прямоугольной декартовой системе координат).

Координатная плоскость, точка (x, y) которой (при любых $x, y \in \boldsymbol{R}$) отождествлена с числом $x + yi$, называется *комплексной плоскостью* и обозначается, как и множество комплексных чисел, буквой $\boldsymbol{C}$.

В плоскости $\boldsymbol{C}$ ось Ox называется *действительной*, а Oy — *мнимой* осью.

Угол φ, образованный радиусом-вектором z, $z \neq 0$, с положительным направлением оси Ox, называется *аргументом* комплексного числа z и обозначается $\operatorname{Arg} z$. Значения φ аргумента комплексного числа z такие, что $-\pi < \varphi \leqslant \pi$, обычно обозначают $\arg z$. Очевидно, что $\operatorname{Arg} z$ определяется комплексным числом $z \neq 0$ с точностью до целочисленного, кратного числу 2π, в то время как $\arg z$ определяется уже числом $z \neq 0$ однозначно. Очевидно также, что

$$\arg z = \operatorname{arctg} \frac{y}{x} + k\pi,$$

где $k = 0$ для первой и четвертой координатных четвертей, $k = 1$ для второй четверти и $k = -1$ для третьей. Если $x = 0$, то при $y \neq 0$ считается, что $\arg z = \frac{\pi}{2}\operatorname{sign} y$, а при $x = y = 0$ $\arg z$ не определен.

Пусть $|z| = r$, $\operatorname{Arg} z = \varphi$; тогда (рис. 98) $x = r\cos\varphi$, $y = r\sin\varphi$, поэтому

$$z = x + yi = r(\cos\varphi + i\sin\varphi).$$

Правая часть этого равенства называется *тригонометрической формой комплексного числа z* в отличие от записи его в виде $z = x + yi$, которая, как было отмечено, называется его алгебраической формой.

Мы будем употреблять тригонометрическую форму записи комплексных чисел и для числа $z = 0$: в этом случае $r = 0$, а φ может принимать любое значение — аргумент нуля не определен.

Комплексные числа $x_1 + y_1 i$ и $x_2 + y_2 i$ считаются равными тогда и только тогда, когда $x_1 = x_2$ и $y_1 = y_2$. По определению полагают также $x + 0i = x$, $0 + yi = yi$, $0 + 0i = 0$.

Сумма двух комплексных чисел $z_1 = x_1 + yi_1$ и $z_2 = x_2 + yi_2$ определяется формулой

$$z_1 + z_2 \overset{\text{def}}{=} (x_1 + x_2) + (y_1 + y_2)i. \tag{19.1}$$

Иначе говоря, действительная и мнимая части суммы $z_1 + z_2$ равны суммам соответственно действительных и мнимых частей z_1 и z_2.

Вычитание комплексных чисел определяется как действие, обратное сложению, т. е. разность $z = z_1 - z_2$ является таким числом z, что $z_2 + z = z_1$. Следовательно, если $z = x + yi$, то $x_2 + x + (y_2 + y)i = x_1 + yi_1$. Отсюда $x = x_1 - x_2$, $y = y_1 - y_2$, т. е. действительная и мнимая части разности $z_1 - z_2$ равны разностям соответственно действительных и мнимых частей чисел z_1 и z_2.

Геометрически действительная и мнимая части комплексного числа при интерпретации его как вектора являются координатами этого вектора, а при сложении (вычитании) координат векторов сами векторы также складываются (вычитаются), поэтому формула (19.1) означает, что геометрически комплексные числа складываются как векторы (рис. 99 и 100).

Модуль комплексного числа $z = x + iy$ является, очевидно, длиной соответствующего ему вектора (x, y). Длина каждой стороны треугольника не превосходит суммы длин двух других его сторон, а абсолютная величина разности длин двух сторон не превосходит длины третьей стороны, поэтому для любых двух комплексных чисел z_1 и z_2 имеют место неравенства (рис. 99 и 100)

$$|z_1 + z_2| \leqslant |z_1| + |z_2|, \quad ||z_1| - |z_2|| \leqslant |z_1 - z_2|.$$

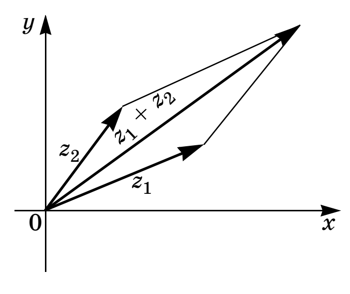


Рис. 99

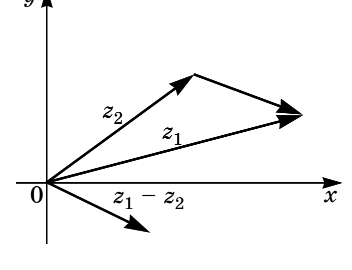


Рис. 100

Второе из этих неравенств можно получить из первого тем же приемом, который был использован для получения аналогичного неравенства для действительных чисел в п. 2.3*.

Произведение двух комплексных чисел $z_1 = x_1 + yi_1$ и $z_2 = x_2 + yi_2$ определяется по формуле

$$z_1 z_2 \equiv (x_1 + yi_1)(x_2 + yi_2) \stackrel{\text{def}}{=} (x_1 x_2 - y_1 y_2) + (x_1 y_2 + y_1 x_2)i. \quad (19.2)$$

Умножим, согласно этому правилу, число $i = 0 + 1i$ само на себя. Получим $ii = -1$. Произведение ii естественно обозначить i^2. Таким образом,

$$i^2 = -1.$$

Если иметь в виду это соотношение, то формула (19.2) означает не что иное, как обычное формальное почленное умножение.

Найдем формулы умножения комплексных чисел в тригонометрической форме. Если

$$z_1 = r_1(\cos \varphi_1 + i\sin \varphi_1),\ z_2 = z_2(\cos \varphi_2 + i\sin \varphi_2),$$

то

$$z_1 z_2 = r_1 r_2[(\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2) +$$
$$+ i(\cos \varphi_1 \sin \varphi_2 + \sin \varphi_1 \cos \varphi_2)] =$$
$$= r_1 r_2[\cos (\varphi_1 + \varphi_2) + i\sin (\varphi_1 + \varphi_2)]$$

и, таким образом,

$$|z_1 z_2| = |z_1|\,|z_2|,\ \operatorname{Arg}(z_1 z_2) = \operatorname{Arg} z_1 + \operatorname{Arg} z_2. \quad (19.3)$$

Второе равенство, как и вообще все равенства, содержащие Arg, следует понимать как равенство соответствующих

множеств, причем каждый элемент множества, стоящего в правой части равенства, является суммой некоторых элементов соответственно из множеств $\mathrm{Arg}\, z_1$ и $\mathrm{Arg}\, z_2$.

Методом математической индукции легко показать, что

$$|z_1 z_2 \ldots z_n| = |z_1|\,|z_2| \ldots |z_n|,$$

$$\mathrm{Arg}\,(z_1 z_2 \ldots z_n) = \mathrm{Arg}\, z_1 + \mathrm{Arg}\, z_2 + \ldots + \mathrm{Arg}\, z_n,$$

т. е. при умножении комплексных чисел их модули перемножаются, а аргументы складываются.

Отсюда, полагая $z_1 = z_2 = \ldots = z_n = z$, для степени z^n, $n = 1, 2, 3, \ldots$, комплексного числа z имеем

$$|z^n| = |z|^n,\ \mathrm{Arg}\, z^n = n\mathrm{Arg}\, z + 2k\pi,\ k = 0, \pm 1, \pm 2, \ldots .$$

Следует обратить внимание на то, что формула для аргумента степени комплексного числа представляет собой равенство множеств: если φ — какое-либо значение аргумента числа z, то множество всех аргументов числа z^n составляют числа вида $n\varphi + 2\pi m$, $m = 0, \pm 1, \pm 2, \ldots$. Отсюда ясно, что если $n \geqslant 2$, то $\mathrm{Arg}\, z^n \neq n\mathrm{Arg}\, z$, так как здесь правая часть представляет собой совокупность всех чисел вида $n(\varphi + 2\pi m) = n\varphi + 2\pi n m$, т. е. к числу $n\varphi$ добавляются не всевозможные, кратные числу 2π, как в случае $\mathrm{Arg}\, z^n$, а лишь кратные числу $2\pi n$.

Отметим еще, что формула для аргумента степени комплексного числа равносильна следующему утверждению.

Если $\varphi \in \mathrm{Arg}\, z$, *то* $n\varphi \in \mathrm{Arg}\, z^n$.

Поэтому если $z = r(\cos\varphi + i \sin\varphi)$, то $z^n = r^n(\cos n\varphi + i \sin n\varphi)$.

Отсюда для комплексного числа, абсолютная величина которого равна 1 ($r = 1$) и, следовательно, имеющего вид $z = \cos\varphi + i\sin\varphi$, имеем

$$(\cos\varphi + i\sin\varphi)^n = \cos n\varphi + i\sin n\varphi. \qquad (19.4)$$

Это соотношение называется *формулой Муавра*[1].

Деление $\frac{z_1}{z_2}$ комплексного числа z_1 на комплексное число $z_2 \neq 0$ определяется как операция, обратная умножению,

[1] А. Муавр (1667—1754) — французский математик.

т. е. число $z = \frac{z_1}{z_2}$ называется *частным от деления* z_1 *на* z_2, *если* $z_1 = z_2 z$. Поэтому

$$|z_1| = |z_2|\,|z| \text{ и } \operatorname{Arg} z_1 = \operatorname{Arg} z_2 + \operatorname{Arg} z,$$

откуда

$$|z| = \left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}, \quad \operatorname{Arg} z = \operatorname{Arg} \frac{z_1}{z_2} = \operatorname{Arg} z_1 - \operatorname{Arg} z_2, \qquad (19.5)$$

т. е. при делении комплексных чисел их модули делятся, а аргументы вычитаются.

Формулами (19.5) комплексное число $z = \frac{z_1}{z_2}$ при заданных z_1 и $z_2 \neq 0$, очевидно, определено однозначно. Ряд других свойств комплексных чисел, как, например, коммутативность и ассоциативность сложения и умножения, дистрибутивность умножения относительно сложения и другие свойства, непосредственно следуют из формул, с помощью которых определены эти операции для комплексных чисел, и из соответствующих свойств действительных чисел. Поэтому не будем на них подробно останавливаться.

Если n — натуральное число, то *корнем* n*-й степени* их комплексного числа $w = \sqrt[n]{z}$ называется такое число w, n-я степень которого равна подкоренному выражению: $w^n = z$.

Например, $i^2 = -1$, поэтому $\sqrt{-1} = i$. Однако этой формулой задаются не все значения корня $\sqrt{-1}$, так как $(-i)^2 = -1$ и, следовательно, $-i$ также корень из $\sqrt{-1}$, т. е. верно не только равенство $\sqrt{-1} = i$, но и равенство $\sqrt{-1} = -i$.

Если

$$z = r(\cos\alpha + i\sin\varphi), \text{ а } w = \rho(\cos\varphi + i\sin\varphi) \text{ и } w^n = z,$$

то

$$\rho^n(\cos n\psi + i\sin n\psi) = r(\cos\varphi + i\sin\varphi);$$

отсюда

$$\rho = \sqrt[n]{r}.$$

Здесь корень понимается в арифметическом смысле — как неотрицательное действительное число, так как, по определению модуля комплексного числа, $\rho \geqslant 0$.

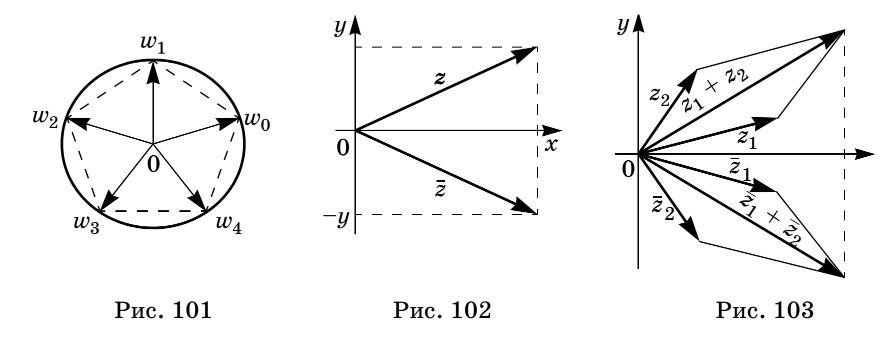


Рис. 101 Рис. 102 Рис. 103

Далее,

$$n\psi = \varphi + 2k\pi \ (k \text{ — целое}), \ \text{или} \ \psi = \frac{\varphi + 2k\pi}{n}.$$

По существу, различные значения аргумента получаются при значениях $k = 0, 1, \dots, n - 1$: различные в том смысле, что если обозначить эти значения аргумента через ψ_k и положить $w_k = \rho(\cos \psi_k + \sin \psi_k)$, то при $\rho \neq 0$ получаются различные комплексные числа. При всех остальных k значения ψ будут отличаться от указанных чисел ψ_k на кратное 2π, т. е. эти значения аргумента будут приводить к одному из комплексных чисел w_k, $k = 0, 1, \dots, n - 1$. Таким образом, корень $\sqrt[n]{z}$ имеет при $z \neq 0$ точно n значений $w_0, w_1, \dots, w_{n-1}$.

В комплексной плоскости числа w_k, $k = 0, 1, \dots, n - 1$, располагаются в вершинах правильного n-угольника, вписанного в круг радиуса ρ с центром в начале координат. Это следует из того, что аргумент числа w_k отличается от аргумента числа w_{k-1} при всех $k = 1, 2, \dots, n - 1$, так же как и аргумент w_{n-1} от аргумента w_0, на одно и то же число $2\pi/n$. На рисунке 101 изображен случай $n = 5$.

Каждому комплексному числу $z = x + yi$ соответствует число $x - yi$, которое называется *сопряженным* с z и обозначается $\bar{z}$, $\bar{z} = x - yi$. Геометрически число $\bar{z}$ изображается вектором, симметричным с вектором z относительно оси Ox (рис. 102).

Свойства сопряженных комплексных чисел

1^0. $|\bar{z}| = |z|, \quad \arg \bar{z} = -\arg z$.

2^0. $z\bar{z} = |z|^2$.

3^0. $z = z$.

4^0. $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$.

5^0. $\overline{z_1 - z_2} = \bar{z}_1 - \bar{z}_2$.

6^0. $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$.

7^0. $\left(\frac{\bar{z}_1}{z_2}\right) = \frac{\bar{z}_1}{\bar{z}_2}$, $z_2 \neq 0$.

Свойство 1^0 очевидно (см. рис. 102). Далее, согласно правилу умножения комплексных чисел,

$$z\bar{z} = (x + yi)(x - yi) = x^2 + y^2 = |z|^2. \ \square$$

Свойство 3^0 также очевидно: если $z = x + yi$, то $\bar{z} = x - yi$ и $\bar{\bar{z}} = \overline{x - yi} = x + yi = z$. □

В справедливости свойства 4^0 можно убедиться геометрически, взяв параллелограмм, симметричный относительно оси Ox с параллелограммом, построенным на векторах z_1 и z_2 как на сторонах (рис. 103), т. е. параллелограмм, натянутый на векторы $\bar{z}_1$ и $\bar{z}_2$. Диагонали этих параллелограммов будут также симметричными друг другу относительно оси Ox и, следовательно, будут соответственно равными $z_1 + z_2$ и $\overline{z_1 + z_2}$. С другой стороны, последняя диагональ, как сумма векторов $\bar{z}_1$ и $\bar{z}_2$, равна также и $\bar{z}_1 + \bar{z}_2$. □

Свойство 5^0 доказывается аналогично.

Свойства 6^0 и 7^0 следуют из того, что модули и аргументы выражений, стоящих в разных частях соответствующих равенств, совпадают. Действительно, используя свойство 1^0, получим

$$|\overline{z_1 z_2}| = |z_1 z_2| = |z_1|\,|z_2| = |\bar{z}_1|\,|\bar{z}_2| = |\bar{z}_1\,\bar{z}_2|,$$

$$\text{Arg}\,\overline{z_1 z_2} = -\,\text{Arg}\,z_1 z_2 = -(\text{Arg}\,z_1 + \text{Arg}\,z_2) =$$

$$= -\text{Arg}\,z_1 - \text{Arg}\,z_2 = \text{Arg}\,\bar{z}_1 + \text{Arg}\,\bar{z}_2 = \text{Arg}\,\bar{z}_1\bar{z}_2. \ \square$$

Аналогично доказывается свойство 7^0.

Используя сопряженные комплексные числа, можно получить формулу для частного комплексных чисел в алгебраической форме: умножив числитель и знаменатель дроби $\frac{x_1 + y_1 i}{x_2 + y_2 i}$ на число $x_2 - y_2 i$, сопряженное знаменателю, получим

$$\frac{x_1 + y_1 i}{x_2 + y_2 i} = \frac{(x_1 + y_1 i)(x_2 - y_2 i)}{(x_2 + y_2 i)(x_2 - y_2 i)} = \frac{x_1 x_2 + y_1 y_2}{x_2^2 + y_2^2} + \frac{y_1 x_2 - x_1 y_2}{x_2^2 + y_2^2} i.$$

19.2*. Формальная теория комплексных чисел

Вдумчивый читатель обратил внимание на то, что приводимая в п. 19.1 формулировка «выражения вида $z = x + iy$ называются комплексными числами» не является четким определением комплексных чисел.

Множество комплексных чисел $\boldsymbol{C}$ можно определить как множество упорядоченных пар (x, y) действительных чисел, $x \in \boldsymbol{R}$, $y \in \boldsymbol{R}$, в котором введены операции сложения и умножения согласно следующему определению:

$$(x, y) + (x', y') \stackrel{\text{def}}{=} (x + x', y + y'),$$

$$(x, y)(x', y') \stackrel{\text{def}}{=} (xx' - yy', xy' + x'y),$$

$$(x, y) \in \boldsymbol{C}, (x', y') \in \boldsymbol{C}.$$

Нетрудно проверить, что в результате этого определения множество указанных пар превращается в поле, т. е. удовлетворяет условиям I, II, III п. 2.1. Полученное таким образом поле, а также каждое изоморфное ему, называется *полем комплексных чисел*.

Пары $(x, 0)$ обозначаются просто через x (их совокупность изоморфна полю действительных чисел), а пара $(0, 1)$ обозначается через $i : i \stackrel{\text{def}}{=} (0, 1)$.

Согласно определенной операции умножения,

$$i^2 = (0, 1)(0, 1) = (-1, 0) = -1, \text{ т. е. } i^2 = -1.$$

Для любого комплексного числа (x, y) имеет место легко проверяемое тождество

$$(x, y) = x + yi.$$

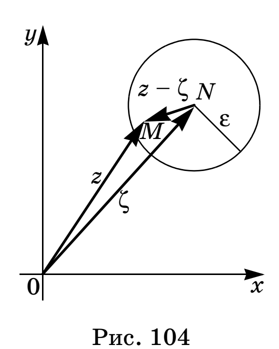


Рис. 104

Действительно,

$$(x, y) = (x, 0) + (0, y) =$$
$$= (x, 0) + (0, 1)(y, 0) = x + yi,$$

и мы снова пришли к записи комплексных чисел, из которой исходили в п. 19.1.

УПРАЖНЕНИЕ 1. Докажите, что матрицы вида $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ с обычными матричными операциями образуют поле, изоморфное полю комплексных чисел (a и b — действительные числа).

19.3. Некоторые понятия анализа в области комплексных чисел

Понятия числовой последовательности и ее предела легко обобщаются и на случай комплексных чисел. Функция, определенная на множестве натуральных чисел и имеющая своими значениями комплексные числа, называется *последовательностью комплексных чисел*. Как и в случае действительных чисел, комплексное число z, соответствующее натуральному числу n, снабжается индексом n: z_n, $n = 1, 2, \ldots$.

Определение 1. *Пусть задана последовательность комплексных чисел $z_n = x_n + y_n i$, $n = 1, 2, \ldots$. Число $\zeta = \xi + \eta i$ называется ее пределом, если для любого действительного числа $\varepsilon > 0$ существует такой номер n_ε, что при $n > n_\varepsilon$ выполняется неравенство $|z_n - \zeta| < \varepsilon$.*

В этом случае пишут $\lim\limits_{n \to \infty} z_n = \zeta$ и говорят, что последовательность $\{z_n\}$ сходится к числу ζ (рис. 104).

Таким образом, по форме это определение совершенно такое же, как для предела последовательности действительных чисел.

Определение 1 означает, что существование предела $\lim\limits_{n \to \infty} z_n = \zeta$ последовательности $\{z_n\}$ комплексных чисел рав-

носильно существованию предела $\lim\limits_{n \to \infty} |z_n - \xi| = 0$ последовательности действительных чисел.

Пусть $\zeta = \xi + \eta i$, $z_n = x_n + y_n i$, $n = 1, 2, \ldots$. Поскольку, очевидно, справедливы неравенства

$$|x_n - \xi| \leqslant |z_n - \zeta|, \ |y_n - \eta| \leqslant |z_n - \zeta|$$

(геометрически: длины катетов не превышают длины гипотенузы) и

$$|z_n - \zeta| \leqslant |x_n - \xi| + |y_n - \eta|$$

(геометрически: длина стороны треугольника не превышает суммы длин двух других его сторон), последовательность $\{z_n = x_n + y_n i\}$ имеет своим пределом число $\zeta = \xi + \eta i$ тогда и только тогда, когда последовательности ее действительных и мнимых частей $\{x_n\}$ и $\{y_n\}$ сходятся соответственно к числам ξ и η. Иначе говоря, равенство $\lim\limits_{n \to \infty} z_n = \zeta$ равносильно равенствам $\lim\limits_{n \to \infty} x_n = \xi$, $\lim\limits_{n \to \infty} y_n = \eta$.

Последовательность комплексных чисел, имеющая своим пределом нуль, называется *бесконечно малой*.

На последовательности комплексных чисел естественным образом переносится ряд теорем о пределах последовательностей действительных чисел, например теорема о единственности предела, об ограниченности последовательности, имеющей предел, критерий Коши и т. п.

В § 8 были введены обозначения o и O для сравнения функций. В дальнейшем понадобятся такие же обозначения и для последовательностей комплексных чисел.

Определение 2. *Будем говорить, что последовательность* $\{z_n\}$ *ограничена относительно последовательности* $\{w_n\}$, *и писать* $z_n = O(w_n)$[1], *если существует постоянная* $c > 0$ *такая, что*

$$|z_n| \leqslant c|w_n|, \ n = 1, 2, \ldots .$$

[1] Иногда к этому добавляют: при $n \to \infty$.

Это определение в случае $w_n \neq 0$, $n = 1, 2, \ldots$, эквивалентно следующему: для двух данных последовательностей $\{z_n\}$ и $\{w_n\}$ существуют постоянная $c' > 0$ и номер n_0 такие, что

$$|z_n| \leqslant c'|w_n|, \quad n = n_0, n_0 + 1, \ldots .$$

Действительно, полагая в этом случае

$$c = \max\left\{\left|\frac{z_1}{w_1}\right|, \left|\frac{z_2}{w_2}\right|, \ldots, \left|\frac{z_{n_0-1}}{w_{n_0-1}}\right|, c'\right\},$$

получаем $|z_n| \leqslant c|w_n|$, $n = 1, 2, \ldots$, т. е. первоначальное определение.

Определение 3. *Если $z_n = O(w_n)$ и $w_n = O(z_n)$, то будем говорить, что последовательности $\{z_n\}$ и $\{w_n\}$ одного порядка и писать $z_n \asymp w_n$.*

Определение 4. *Будем говорить, что последовательность $\{z_n\}$ является бесконечно малой по сравнению с последовательностью $\{w_n\}$, и писать $z_n = o(w_n)$, если существует бесконечная малая последовательность $\{\alpha_n\}$ такая, что $z_n = \alpha_n w_n$, $n = 1, 2, \ldots$.*

Определение 5. *Последовательности $\{z_n\}$ и $\{w_n\}$ называются эквивалентными или асимптотически равными, если существует такая последовательность $\{\varepsilon_n\}$, что*

$$\lim_{n\to\infty} \varepsilon_n = 1 \quad \text{и} \quad z_n = \varepsilon_n w_n, \quad n = 1, 2, \ldots .$$

В этом случае пишут $z_n \sim w_n$, $n = 1, 2, \ldots$.

УПРАЖНЕНИЯ. 2. Докажите, что, для того чтобы $z_n \sim w_n$, необходимо и достаточно, чтобы $z_n = w_n + o(w_n)$, $n = 1, 2, \ldots$.

3. Докажите: если $z_n = cw_n + o(w_n)$, $n = 1, 2, \ldots$, то $z_n = O(w_n)$.

Можно рассматривать и функции комплексного аргумента. Например, $f(z) = |z|$, $f(z) = z^2$. Обе эти функции определены на множестве всех комплексных чисел, первая из них принимает только неотрицательные действительные значения, вторая же, наряду с действительными, и существенно комплексные.

Если функция $w = f(z)$ задана на множестве $\boldsymbol{Z} \subset \boldsymbol{C}$, $z \in \boldsymbol{Z}$, $w \in \boldsymbol{C}$, то геометрически функция f задает отображение пло-

ского множества Z в плоскость комплексной переменной w. Например, функция $w = |z|$ отображает плоскость на полупрямую, а функция $w = z^2$ — всю плоскость на всю плоскость, как говорят, двукратным образом — в данном случае это означает, что при отображении $w = z^2$ каждая точка образа кроме нуля имеет прообраз, состоящий из двух точек.

Для комплекснозначных функций можно ввести многие из понятий, введенных ранее для действительнозначных функций (предел, непрерывность, дифференцируемость, интеграл и др.). В ближайших параграфах придется встретиться лишь с понятием ограниченности и непрерывности комплекснозначных функций.

Комплекснозначная функция $f(z)$, $z \in \boldsymbol{Z} \subset \boldsymbol{C}$ называется *ограниченной на множестве* $\boldsymbol{Z}$, если на этом множестве ограничена функция $|f(z)|$.

Таким образом, понятие ограниченности комплекснозначной функции f сводится к понятию ограниченности действительнозначной функции $|f|$.

Определение 6. *Пусть комплекснозначная функция f определена на множестве $\boldsymbol{Z} \subset \boldsymbol{C}$ и пусть $z_0 \in \boldsymbol{C}$. Функция f называется непрерывной в точке z_0, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для всех точек $z \in \boldsymbol{Z}$, удовлетворяющих условию $|z - z_0| < \delta$, выполняется неравенство $|f(z) - f(z_0)| < \varepsilon$.*

Мы видим, что по форме это определение полностью совпадает с определением непрерывности для действительных функций (см. п. 5.5).

Комплекснозначная функция, непрерывная в каждой точке некоторого множества, называется непрерывной на этом множестве. Из определения непрерывности функции и неравенства $|f(z)| - |f(z_0)| \leqslant |f(z) - f(z_0)|$ следует, что если функция $f(z)$, определенная на множестве $\boldsymbol{Z} \subset \boldsymbol{C}$, непрерывна в какой-то точке z_0 этого множества: $z_0 \in \boldsymbol{Z}$, то и действительнозначная функция $|f(z)|$ непрерывна в этой точке.

На комплекснозначные функции переносится теорема о том, что если две функции f и g, определенные на некотором множестве $\boldsymbol{Z} \subset \boldsymbol{C}$, непрерывны в точке $z_0 \in \boldsymbol{Z}$, то и функции $f + g$, fg, а если $g(z_0) \neq 0$, то и f/g, непрерывны в этой точке.

Из этой теоремы следует, например, что любой многочлен $P_n(z) = \sum_{k=0}^{n} a_k z^k$ с комплексными коэффициентами a_k, $k = 0, 1, \ldots, n$, непрерывен в любой точке $z_0 \in \boldsymbol{C}$ (ср. с п. 7.1).

19.4. Разложение многочленов на множители

Пусть

$$P_n(z) = A_n z^n + A_{n-1} z^{n-1} + \ldots + A_1 z + A_0,\ z \in \boldsymbol{C}, \qquad (19.6)$$

— многочлен с комплексными в общем случае коэффициентами A_j, $j = 0, 1, \ldots, n$.

Как и в случае многочленов с действительными коэффициентами (см. п. 5.3) при $A_n \neq 0$ число n называется *степенью многочлена*, а нулевому многочлену приписывается степень $-\infty$.

Из алгебры известно, что если степень m многочлена $Q_m(z)$ не превышает степени n многочлена $P_n(z)$, то существуют такие многочлен $S_k(z)$ степени k и многочлен $R_l(z)$ степени l или равный нулю, что $n = m + k$, $0 \leqslant l < m$, и многочлен $P_n(z)$ представим в виде

$$P_n(z) = S_k(z) Q_m(z) + P_l(z).$$

При этом такое представление единственно.

Операция нахождения многочленов $S_k(z)$ и $P_l(z)$ по заданным многочленам $P_n(z)$ и $Q_m(z)$ называется *делением многочлена $P_n(z)$ на $Q_m(z)$*, многочлен $P_m(z)$ — *делимым*, $Q_m(z)$ — *делителем*, S_k — *частным*, $R_l(z)$ — *остатком* от деления $P_n(z)$ на $Q_m(z)$.

Отметим, что из $m = 1$ следует, что $l = 0$, т. е. в этом случае остаток от деления является константой.

Комплексное число z_0 такое, что $P_n(z_0) = 0$, называется *корнем* данного многочлена (19.6).

Если многочлен $P_n(z)$ степени $n \geqslant 1$ разделить на $z - \zeta$, где ζ — какое-либо комплексное число, то получим

$$P_n(z) = (z - \zeta) Q_{n-1}(z) + r,$$

где $Q_{n-1}(z)$ — многочлен степени $n - 1$, а остаток r — постоянная. Отсюда непосредственно следует, что число z_0 является корнем многочлена $P_n(z)$ тогда и только тогда, когда

многочлен $P_n(z)$ делится без остатка на $z - z_0$ (теорема Безу[1]).

Если многочлен $P_n(z)$ делится на $(z - z_0)^k$ (k — положительное целое) и не делится на $(z - z_0)^{k+1}$, то число k называется *кратностью корня* z_0.

Однократный корень называется *простым*, а k-кратный при $k > 1$ — *кратным корнем*.

Таким образом, если комплексное число z_0 является корнем кратности k многочлена $P_n(z)$, то

$$P_n(z) = (z - z_0)^k Q_{n-k}(z),$$

где $Q_{n-k}(z)$ — такой многочлен степени $n - k$, что $Q_{n-k}(z_0) \neq 0$.

В курсе алгебры доказывается, что *всякий многочлен* $P_n(z)$ *степени* $n \geqslant 1$ *имеет по крайней мере один корень* z_1. Если его кратность равна k, то, как отмечалось, справедливо разложение.

$$P_n(z) = (z - z_1)^{k_1} Q_{n-k_1}(z), \quad Q_{n-k_1}(z_1) \neq 0,$$

где степень многочлена $Q_{n-k_1}(z)$ меньше n. Многочлен $Q_{n-k_1}(z)$, если его степень больше 1, также имеет хотя бы один корень z_2. Если кратность этого корня равна k_2, то

$$P_n(z) = (z - z_1)^{k_1} (z - z_2)^{k_2} Q_{n-k_1-k_2}(z),$$

$$Q_{n-k_1-k_2}(z_1) \neq 0, \; Q_{n-k_1-k_2}(z_2) \neq 0.$$

Продолжая этот процесс дальше, через конечное число m шагов получим многочлен нулевой степени $P_{n-k_1-\ldots-k_m}(z) = A_n$ и, следовательно, для многочлена $P_n(z)$ справедливо следующее разложение на множители:

$$P_n(z) = A_n(z - z_1)^{k_1} (z - z_2)^{k_2} \ldots (z - z_m)^{k_m}, \qquad (19.7)$$

где $k_1 + k_2 + \ldots + k_m = n$, откуда следует, что *каждый многочлен степени* $n \geqslant 1$ *имеет точно* n *корней, если каждый корень считать столько раз, какова его кратность.*

[1] Э. Безу (1730—1783) — французский математик.

Для многочлена (19.6) обозначим через $\overline{P}_n(z)$ многочлен, коэффициенты которого являются комплексными числами, сопряженными коэффициентам многочлена $P_n(z)$:

$$\overline{P}_n(z) = \overline{A}_n z^n + \overline{A}_{n-1} z^{n-1} + \ldots + \overline{A}_1 z + \overline{A}_0.$$

Многочлен $\overline{P}_n(z)$ называется многочленом, *сопряженным* многочлену $P_n(z)$.

В силу свойств сопряженных комплексных чисел, имеем

$$\overline{P_n(z)} = \overline{P}_n(\bar{z}).$$

Действительно,

$$\overline{P_n(z)} = \overline{A_n z^n + A_{n-1} z^{n-1} + \ldots + A_n z + A_0} =$$

$$= \overline{A}_n \bar{z}^n + \overline{A}_{n-1} \bar{z}^{n-1} + \ldots + \overline{A}_1 \bar{z} + \overline{A}_0 = \overline{P}_n(\bar{z}).$$

Очевидно также, что $\bar{p}_n(z) = P(z)$.

Покажем, что если число z_0 является корнем многочлена $P_n(z)$ кратности k, то сопряженное ему число $\bar{z}_0$ является корнем сопряженного многочлена $\overline{P}_n(z)$ и притом той же кратности.

В самом деле, переходя в формулах

$$P_n(z) = (z - z_0)^k Q_{n-k}(z), \;\; Q_{n-k}(z_0) \neq 0,$$

к сопряженным выражениям, получим

$$\overline{P}_n(\bar{z}) = (\bar{z} - \bar{z}_0)^k \overline{Q}_{n-k}(\bar{z}), \;\; \overline{Q}_{n-k}(\bar{z}_0) \neq 0.$$

Полагая для наглядности $\zeta = \bar{z}$ ($\bar{z}$, как и z — произвольные комплексные числа), перепишем полученные формулы в виде

$$\overline{P}_n(\zeta) = (\zeta - \bar{z}_0)^k Q_{n-k}(\zeta), \;\; \overline{Q}_{n-k}(\bar{z}_0) \neq 0.$$

Это и означает, что число $\bar{z}_0$ является корнем кратности k для многочлена $\overline{P}_n(z)$.

Пусть теперь все коэффициенты многочлена $P_n(z)$ суть действительные числа. В этом случае сопряженный многочлен $\overline{P}_n(z)$, очевидно, совпадает с самим многочленом $P_n(z)$. Поэтому из доказанного следует, что если комплексное число z_0 является корнем кратности k многочлена $P_n(z)$ с действительными коэффициентами, то и сопряженное ему число $\bar{z}_0$ также является корнем кратности k этого многочлена.

Отметим далее, что произведение $(z - z_0)(z - \bar{z}_0)$ всегда является многочленом (относительно z) с действительными коэффициентами. Действительно, пусть $z_0 = a + bi$, где a и b действительны. Тогда $\bar{z}_0 = a - bi$, поэтому

$$(z - z_0)(z - \bar{z}_0) = (z - a - bi)(z - a + bi) = (z - a)^2 + b^2 =$$
$$= z^2 - 2az + a^2 + b^2 = z^2 + pz + q, \qquad (19.8)$$

где положено $p = -2a$ и $q = a^2 + b^2$, очевидно, p и q действительны. Отметим, что $p^2/4 - q = -b^2$, поэтому при $b \neq 0$, т. е. тогда, когда корень z_0 является существенно комплексным числом, выполняется неравенство

$$\frac{p^2}{4} - q < 0. \qquad (19.9)$$

Обратим внимание и на справедливость обратного утверждения: если выполнено неравенство (19.9), то корни трехчлена $z^2 + pz + q$ (p и q действительны) — существенно комплексные числа.

Из сказанного следует, что для всякого многочлена степени n с действительными коэффициентами справедливо разложение на множители вида

$$P_n(x) = A_n(x - a_1)^{\alpha_1} \ldots (x - a_r)^{\alpha_r}(x^2 + p_1x + q_1)^{\beta_1} \ldots$$
$$\ldots (x^2 + p_sx + q_s)^{\beta_s}, \qquad (19.10)$$

где

$$\sum_{i=0}^{r}\alpha_i + 2\sum_{i=0}^{s}\beta_j = n, \quad \frac{p_j^2}{4} - q_j < 0, \quad j = 1, 2, \ldots, s,$$

и все коэффициенты $A_n, a_1, \ldots, a_r, p_1, q_1, \ldots, p_s, q_s$ действительны. При этом $a_1, \ldots, a_r$ — все действительные корни многочлена $P_n(x)$, а каждому существенно комплексному корню z_0 и ему сопряженному корню $\bar{z}_0$ соответствует множитель вида $x^2 + px + q = (x - z_0)(x - \bar{z}_0)$. Вместо буквы z, употреблявшейся выше для обозначения аргумента рассматриваемого многочлена, здесь по традиции написана буква x, чтобы подчеркнуть, что все рассмотрения происходят в действительной области.

Формула (19.10) непосредственно следует из формул (19.7) и (19.8): нужно в разложении (19.7) сгруппировать попарно множители с сопряженными корнями и записать произведения вида $(z - z_0)(z - \bar{z}_0)$ в форме (19.8). Тогда, замечая, что

кратность сопряженных корней z_0 и $\bar{z}_0$ одинакова, мы и получим формулу (19.10).

Разложение многочлена на множители вида (19.10) единственно, так как оно однозначно определяется корнями этого многочлена и их кратностями.

УПРАЖНЕНИЕ 4. Докажите, что если существует $n + 1$ разных чисел $z_1, z_2, \ldots, z_{n+1}$, в которых многочлен $P(z)$ степени, не превосходящей n, обращается в нуль, то все его коэффициенты равны нулю.

Указание. Воспользоваться тем, что определитель Вандермонда

$$\begin{vmatrix} 1 & 1 & \ldots & 1 \\ z_1 & z_2 & \ldots & z_{n+1} \\ \ldots & \ldots & \ldots & \ldots \\ z_1^n & z_2^n & \ldots & z_{n+1}^n \end{vmatrix}$$

в данном случае не равен нулю.

19.5*. Наибольший общий делитель многочленов

Пусть дан многочлен $P(x)$. Всякий многочлен $R(x)$, на который делится многочлен $P(x)$, т. е.

$$P(x) = R(x)r(x), \tag{19.11}$$

где $r(x)$ — также многочлен, называется *делителем* многочлена $P(x)$.

Было показано, что многочлен $P(x)$ можно записать в виде

$$P(x) = A(x - a_1)^{\alpha_1} \ldots (x - a_r)^{\alpha_r} (x^2 + p_1 x + q_1)^{\beta_1} \ldots$$
$$\ldots (x^2 + p_s x + q_s)^{\beta_s}, \tag{19.12}$$

где $a_1, \ldots, a_r$ — действительные корни многочлена, а множители вида $x^2 + p_j x + q_j$ соответствуют существенно комплексным корням этого многочлена,

$$\frac{p_j^2}{4} - q_j < 0, \ \ j = 1, 2, \ldots, s;$$

коэффициенты A, p_j и q_j ($j = 1, 2, \ldots, s$) действительны. Отсюда следует, что всякий делитель $R(x)$ многочлена $P(x)$ может быть записан в виде

$$R(x) = B(x - a_1)^{\lambda_1} \ldots (x - a_r)^{\lambda_r} (x^2 + p_1 x + q_1)^{\mu_1} \ldots$$
$$\ldots (x^2 + p_s x + q_s)^{\mu_s}, \tag{19.13}$$

где

$$\lambda_i \leqslant \alpha_i, \quad i = 1, 2, \ldots, \quad r, \quad \mu_i \leqslant \beta_j, \quad j = 1, 2, \ldots, s. \quad (19.14)$$

Действительно, никаких других множителей вида

$$x - a \quad \text{и} \quad x^2 + px + q, \quad (19.15)$$

где a, p и q действительны и $p^2/4 - q < 0$ в разложении многочлена $R(x)$ быть не может, ибо, с одной стороны, многочлен $R(x)$, как всякий многочлен, может быть разложен на множители вида (19.15), с другой стороны, из формулы (19.11) следует, что если в разложении $R(x)$ на множители имеется множитель вида $x - a$, соответственно вида $x^2 + px + q$, то $x = a$, соответственно корни трехчлена $x^2 + px + q$, являются и корнями многочлена $P(x)$; поэтому указанные множители входят в разложение (19.12). Неравенства (19.14) также очевидны: из той же формулы (19.11) следует, что кратность корня многочлена $R(x)$ не может превышать кратности того же корня многочлена $P(x)$.

Пусть теперь даны два многочлена $P(x)$ и $Q(x)$. Всякий многочлен, являющийся делителем как многочлена $P(x)$, так и многочлена $Q(x)$, называется их *общим делителем*. Общий делитель двух многочленов, который делится на любой общий делитель этих многочленов, называется их *наибольшим общим делителем*.

Если многочлены $P(x)$ и $Q(x)$ записаны в виде (19.12):

$$P(x) = A'(x - a_1')^{\alpha_1'} \ldots (x - \alpha_{r'})^{\alpha_{r'}}(x^2 + p_1' + q_1')^{\beta_1'} \ldots (x^2 + p_s'x + q_{s'}')^{\beta_{s'}'}, \quad (19.16)$$

$$Q(x) = A''(x - a_1'')^{\alpha_1''} \ldots (x - a_{r''}'')^{\alpha_{r''}''}(x^2 + p_1''x + q_1'')^{\beta_1''} \ldots (x^2 + p_{s''}''x + q_{s''}'')^{\beta_{s''}''}, \quad (19.17)$$

то всякий их общий делитель $R(x)$ можно записать в виде (19.13), где множители

$$x - a_k \quad (k = 1, 2, \ldots, r),$$

$$x^2 + p_l x + q_l \quad (l = 1, 2, \ldots, s) \quad (19.18)$$

входят в разложение (19.16) и в разложение (19.17).

Пусть индексы у коэффициентов множителей (19.18) в разложениях (19.16) и (19.17) равны соответственно i'_k, j'_l и i''_k, j''_l, тогда, в силу неравенств (19.14), имеем

$$\lambda_k \leqslant \alpha'_{i'_k},\ \lambda_k \leqslant \alpha''_{i''_k},\ k = 1, 2, \dots , r, \qquad (19.19)$$

$$\mu_l \leqslant \beta''_{j''_k},\ \mu_l \leqslant \beta''_{j''_t},\ l = 1, 2, \dots , s.$$

Для того чтобы многочлен (19.13) был наибольшим общим делителем многочленов $P(x)$ и $Q(x)$, необходимо и достаточно, чтобы показатели степени λ_k, $k = 1, 2, \dots , r$ и μ_l, $l = 1, 2, \dots , s$ были максимальными из возможных, т. е. чтобы

$$\lambda_k = \min \{\alpha'_{i'_k}, \alpha''_{i''_k}\},\ k = 1, 2, \dots , r, \qquad (19.20)$$
$$\mu_l = \min \{\beta'_{i'_l}, \beta'_{j''_l}\},\ l = 1, 2, \dots , s.$$

Действительно, при выполнении этих условий многочлен $R(x)$ будет общим делителем многочленов $P(x)$ и $Q(x)$, кроме того, он будет делиться на любой многочлен вида (19.13), для которого выполнены условия (19.19), т. е. $R(x)$ будет делиться на любой общий делитель многочленов $P(x)$ и $Q(x)$. □

Из найденного вида общего делителя и, в частности, наибольшего общего делителя следует, во-первых, что наибольший делитель двух многочленов не единствен; однако два наибольших общих делителя двух данных многочленов могут отличаться друг от друга лишь постоянным множителем (постоянную B в формуле (19.13) можно брать произвольной, неравной нулю); во-вторых, что наибольший общий делитель двух многочленов имеет степень, бо́льшую, чем любой их общий делитель, не являющийся наибольшим общим делителем.

В качестве примера, полезного для дальнейшего, найдем наибольший общий делитель многочлена $P(x)$ и его производной $P'(x)$. Предварительно заметим, что если число a является действительным корнем кратности α многочлена $P(x)$, т. е.

$$P(x) = (x - a)^{\alpha} P_1(x),\ P_1(a) \neq 0, \qquad (19.21)$$

то a является корнем кратности $\alpha - 1$ для многочлена $P'(x)$.

Действительно, дифференцируя (19.21), имеем

$$P'(x) = \alpha(x - a)^{\alpha - 1} P_1(x) + (x - a)^{\alpha} P'_1(x) = (x - a)^{\alpha - 1} P_2(x),$$

где $P_2(x) = \alpha P_1(x) + (x - a) P'_1(x)$ и $P_2(a) = \alpha P_1(a) \neq 0$.

Подобным образом, если

$$P(x) = (x^2 + px + q)^{\beta} P_3(x), \qquad (19.22)$$

где $p^2/4 - q < 0$, и, значит, корни z_1 и z_2 ($z_2 = \bar{z}_1$) трехчлена $x^2 + px + q$ существенно комплексны, и если

$$P_3(z_1) \neq 0,\ P_3(z_2) \neq 0, \text{ то } P'(x) = (x^2 + px + q)^{\beta - 1} P_4(x),$$

где $P_4(z_1) \neq 0$, $P_4(z_2) \neq 0$, т. е. $P_4(z)$ не делится на $x^2 + px + q$.

Действительно, дифференцируя (19.22), получаем

$$P'(x) = \beta(x^2 + px + q)^{\beta - 1}(2x + p)P_3(x) + (x^2 + px + q)^{\beta} P_3'(x) =$$
$$= (x^2 + px + q)^{\beta - 1} P_4(x),$$

где

$$P_4(x) = \beta(2x + p)P_3(x) + (x^2 + px + q) P_3'(x),$$

откуда следует, что

$$P_4(z_1) = \beta(2z_1 + p)P_3(z_1) \neq 0, \quad P_4(z_2) = \beta(2z_2 + p)P_3(z_2) \neq 0,$$

ибо $z_1 \neq -p/2$ и $z_2 \neq -p/2$, так как они существенно комплексны. □

Из доказанного следует, что если многочлен $P(x)$ записан в виде (19.12), то его производную $P'(x)$ можно представить в виде

$$P'(x) = C(x - a_1)^{\alpha_1 - 1} \ldots (x - a_r)^{\alpha_r - 1} (x^2 + p_1 x + q_1)^{\beta_1 - 1} \ldots$$
$$\ldots (x^2 + p_s x + q_s)^{\beta_s - 1} P_5(x),$$

где многочлен $P_5(x)$ не делится ни на $x - a_i$, $i = 1, 2, \ldots, r$, ни на $x^2 + p_j x + q_j$, $j = 1, 2, \ldots, s$, т. е. не имеет общих корней с многочленом $P(x)$.

Из формул (19.13) и (19.20) получаем, что наибольший общий делитель $R(x)$ многочлена $P(x)$ и его производной $P'(x)$ имеет вид

$$R(x) = (x - a_1)^{\alpha_1 - 1} \ldots (x - a_r)^{\alpha_r - 1} (x^2 + p_1 x + q_1)^{\beta_1 - 1} \ldots$$
$$\ldots (x^2 + p_s x + q_s)^{\beta_s - 1}. \qquad (19.23)$$

Изложенный выше метод получения наибольшего общего делителя двух многочленов $P(x)$ и $Q(x)$ принципиально полностью решает вопрос о существовании и виде наибольшего общего делителя. Практическое же его применение может, однако, быть достаточно трудным: для использования этого метода надо знать разложения на множители вида (19.16) и (19.17) данных многочленов $P(x)$ и $Q(x)$, которые не всегда удается написать в явном виде.

Существует, однако, другой способ получения наибольшего общего делителя двух многочленов $P(x)$ и $Q(x)$, называемый обычно *алгоритмом Евклида*[1]. Опишем его.

Пусть для определенности степень многочлена $P(x)$ больше или равна степени многочлена $Q(x)$. Разделив $P(x)$ на $Q(x)$, получим в качестве частного некоторый многочлен $Q_1(x)$ и остаток $R_1(x)$, который является либо нулем, либо его степень меньше степени многочлена $Q(x)$ (в противном случае процесс деления на $Q(x)$ можно было бы продолжить и дальше):

$$P(x) = Q(x)Q_1(x) + R_1(x).$$

Из этой формулы следует: 1) если многочлены $P(x)$ и $Q(x)$ делятся на некоторый многочлен $R(x)$, то многочлен $R_1(x)$ делится на этот многочлен; 2) если многочлены $Q(x)$ и $R_1(x)$ делятся на какой-то многочлен $R(x)$, то и многочлен $P(x)$ делится на этот многочлен $R(x)$. Отсюда, в свою очередь, следует, что общие делители многочленов $P(x)$ и $Q(x)$, в частности их наибольшие общие делители, совпадают с общими делителями, соответственно с наибольшими общими делителями, многочленов $Q(x)$ и $R_1(x)$.

Разделим далее многочлен $Q(x)$ на многочлен $R_1(x)$:

$$Q(x) = R_1(x)Q_2(x) + R_2(x);$$

продолжая процесс далее, будем иметь

$$R_1(x) = R_2(x)Q_3(x) + R_3(x),$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$R_{k-2}(x) = R_{k-1}(x)Q_k(x) + R_k(x).$$

---

[1] Евклид (ок. 365 г. — ок. 300 г. до н. э.) — древнегреческий математик.

Степени многочленов $R_i(x)$, $i = 1, 2, \ldots$, строго убывают, поэтому существует номер (обозначим его $m + 1$) такой, что $R_{m+1}(x) = 0$, следовательно, $R_{m-1}(x) = R_m(x)Q_{m+1}(x)$.

Пары многочленов $P(x)$ и $Q(x)$, $Q(x)$ и $R_1(x)$, $R_1(x)$ и $R_2(x)$, ... , $R_{m-1}(x)$ и $R_m(x)$ имеют одинаковые общие делители, а значит, и одинаковые наибольшие общие делители. Но $R_{m-1}(x)$ делится на $R_m(x)$, поэтому $R_m(x)$ является наибольшим общим делителем $R_{m-1}(x)$ и $R_m(x)$, следовательно, и наибольшим общим делителем многочленов $P(x)$ и $Q(x)$.

19.6. Разложение правильных рациональных дробей на элементарные

Пусть $P(x)$ и $Q(x)$ — многочлены, вообще говоря, с комплексными коэффициентами, причем $Q(x)$ — ненулевой многочлен.

Рациональная дробь $\frac{P(x)}{Q(x)}$ *называется правильной, если степень многочлена* $P(x)$ *меньше степени многочлена* $Q(x)$, *и неправильной, если степень многочлена* $P(x)$ *не меньше степени многочлена* $Q(x)$.

Всякая рациональная дробь является либо правильной, либо неправильной.

Если рациональная дробь $\frac{P(x)}{Q(x)}$ неправильная, то, разделив числитель на знаменатель по правилу деления многочленов, получим равенство

$$\frac{P(x)}{Q(x)} = R(x) + \frac{P_1(x)}{Q_1(x)}, \qquad (19.24)$$

где $R(x)$, $P_1(x)$ и $Q_1(x)$ — некоторые многочлены, а $\frac{P_1(x)}{Q_1(x)}$ — правильная рациональная дробь.

ЛЕММА 1. *Пусть* $\frac{P(x)}{Q(x)}$ *— правильная рациональная дробь,* $P(x)$ *и* $Q(x)$ *— многочлены с действительными коэффициентами. Если число* a *является действительным корнем кратности* $\alpha \geqslant 1$ *многочлена* $Q(x)$, *т. е.*

$$Q(x) = (x - a)^{\alpha} Q_1(x) \text{ и } Q_1(a) \neq 0,$$

то существуют действительное число A и многочлен $P_1(x)$ с действительными коэффициентами такие, что

$$\frac{P(x)}{Q(x)} = \frac{A}{(x-a)^{\alpha}} + \frac{P_1(x)}{(x-a)^{\alpha-1}Q_1(x)},$$

где дробь $\dfrac{P_1(x)}{(x-a)^{\alpha-1}Q_1(x)}$ *также является правильной.*

Доказательство. Каково бы ни было действительное число A, вычитая из дроби

$$\frac{P(x)}{Q(x)} = \frac{P_1(x)}{(x-a)^{\alpha-1}Q_1(x)}$$

выражение $\dfrac{A}{(x-a)^{\alpha}}$ и затем прибавляя его, получим

$$\frac{P(x)}{Q(x)} = \frac{A}{(x-a)^{\alpha}} + \left[\frac{P_1(x)}{(x-a)^{\alpha-1}Q_1(x)} - \frac{A}{(x-a)^{\alpha}}\right] =$$

$$= \frac{A}{(x-a)^{\alpha}} + \frac{P(x)-AQ_1(x)}{(x-a)^{\alpha}Q_1(x)}. \quad (19.25)$$

По условию, степень многочлена $P(x)$ меньше степени многочлена $Q(x) = (x-a)^{\alpha}Q_1(x)$. Очевидно, что и степень многочлена $Q_1(x)$ меньше степени многочлена $Q(x)$ (так как $\alpha \geqslant 1$), поэтому при любом выборе числа A рациональная дробь $\dfrac{P(x)-AQ_1(x)}{(x-a)^{\alpha}Q_1(x)} = \dfrac{P(x)-AQ_1(x)}{Q(x)}$ является правильной.

Выберем теперь число A таким образом, чтобы число a было корнем многочлена $P(x) - AQ_1(x)$ и, следовательно, чтобы этот многочлен делился на $x - a$. Иначе говоря, определим A из условия $P(a) - AQ_1(a) = 0$; поскольку, по условию, $Q_1(a) \neq 0$, отсюда имеем $A = \dfrac{P(a)}{Q_1(a)}$.

При таком и только таком выборе числа A дробь $\dfrac{P(x)-AQ_1(x)}{(x-a)^{\alpha}Q_1(x)}$ сократится на $x - a$. В результате в этом и только в этом случае после сокращения указанной дроби получится дробь вида $\dfrac{P_1(x)}{(x-a)^{\alpha-1}Q_1(x)}$. Эта дробь получена со-

кращением правильной рациональной дроби с действительными коэффициентами на множитель $x - a$, где a действительно, поэтому и сама она является также правильной рациональной дробью с действительными коэффициентами.

В результате получается искомое разложение, в котором коэффициент A однозначно определен. □

Замечание 1. Если многочлены $P(z)$ и $Q(z)$ имеют комплексные коэффициенты и $z = a$ — комплексный корень кратности $\alpha \geqslant 1$ многочлена $Q(x)$, то разложение (19.25) также имеет место, но число A в этом случае является, вообще говоря, уже комплексным числом. Справедливость этого непосредственно следует из рассуждений, проведенных при доказательстве леммы 1.

ЛЕММА 2. *Пусть* $\frac{P(x)}{Q(x)}$ *— правильная рациональная дробь,* $P(x)$ *и* $Q(x)$ *— многочлены с действительными коэффициентами. Если комплексное число* $z_1 = a + bi$ *(*a *и* b *действительные,* $b \neq 0$*) является корнем кратности* $\beta \geqslant 1$ *многочлена* $Q(x)$*, т. е.*

$$Q(x) = (x^2 + px + q)^{\beta} Q_1(x),$$

где $Q_1(z_1) \neq 0$*, а* $x^2 + px + q = (x - z_1)(x - \bar{z}_1)$*, то существуют действительные числа* M*,* N *и многочлен* $P_1(x)$ *с действительными коэффициентами такие, что*

$$\frac{P(x)}{Q(x)} = \frac{Mx + N}{(x^2 + px + q)^{\beta}} + \frac{P_1(x)}{(x^2 + px + q)^{\beta - 1} Q_1(x)},$$

где дробь $\frac{P_1(x)}{(x^2 + px + q)^{\beta - 1} Q_1(x)}$ *также является правильной.*

Доказательство. Для любых действительных M и N

$$\frac{P(x)}{Q(x)} = \frac{P_1(x)}{(x^2 + px + q)^{\beta} Q_1(x)} =$$

$$= \frac{Mx + N}{(x^2 + px + q)^{\beta}} + \left[\frac{P(x)}{(x^2 + px + q)^{\beta} Q_1(x)} - \frac{Mx + N}{(x^2 + px + q)^{\beta}}\right] =$$

$$= \frac{Mx + N}{(x^2 + px + q)^{\beta}} + \frac{P(x) - (Mx + N)Q_1(x)}{(x^2 + px + q)^{\beta - 1} Q_1(x)}, \qquad (19.26)$$

причем второе слагаемое правой части равенства (19.26) является, как нетрудно видеть, правильной дробью.

Постараемся теперь подобрать M и N так, чтобы числитель этой дроби делился на $x^2 + px + q = (x - z_1)(x - \bar{z}_1)$. Для этого необходимо и достаточно выбрать M и N так, чтобы z_1 было корнем многочлена $P(x) - (Mx + N)Q_1(x)$. Действительно, тогда, согласно сказанному в п. 19.3, число $\bar{z}_1$, сопряженное с z_1, также будет являться корнем указанного многочлена. Отсюда и следует, что этот многочлен в силу существования его разложения вида (19.10) делится на $x^2 + px + q$. Итак, пусть

$$P(z_1) - (Mz_1 + N)Q_1(z_1) = 0. \qquad (19.27)$$

Если это имеет место, то $Mz_1 + N = \frac{P(z_1)}{Q_1(z_1)}$, где, по условию, $Q_1(z_1) \neq 0$.

Пусть $z_1 = a + bi$, $P(z_1)/Q_1(z_1) = A + Bi$; тогда $Mz_1 + N = A + Bi$, или

$$M(a + bi) + N = A + Bi.$$

Отсюда, приравнивая действительные и мнимые части, получим уравнения $Ma + N = A$ и $Mb = B$ и, следовательно, коэффициенты M и N однозначно определяются по формулам

$$M = \frac{B}{b} \quad \text{и} \quad N = A - \frac{a}{b}B. \qquad (19.28)$$

При этих значениях M и N многочлен $P(x) - (Mx + N)Q_1(x)$ делится на многочлен $x^2 + px + q$. Сокращая второе слагаемое правой части равенства (19.26) на $x^2 + px + q$, получим дробь вида

$$\frac{P_1(x)}{(x^2 + px + q)^{\beta - 1}Q_1(x)}.$$

Эта дробь получена сокращением правильной рациональной дроби с действительными коэффициентами на многочлен с действительными коэффициентами, поэтому и сама она является также правильной рациональной дробью с действительными коэффициентами. □

Сформулируем теперь основную теорему этого пункта.

ТЕОРЕМА 1. *Пусть $P(x)/Q(x)$ — правильная рациональная дробь*[1], *$P(x)$ и $Q(x)$ — многочлены с действительными коэффициентами. Если*

$$Q(x) = (x - a_1)^{\alpha_1} \ldots (x - a_r)^{\alpha_r}(x^2 + p_1 x + q_1)^{\beta_1} \ldots (x^2 + p_s x + q_s)^{\beta_s}, \tag{19.29}$$

a_i — попарно различные действительные корни многочлена $Q(x)$ кратности α_i, $i = 1, 2, \ldots, r$,

$$x^2 + p_j x + q_j = (x - z_j)(x - \bar{z}_j),\ z_j \text{ и } \bar{z}_j$$

— попарно различные при разных j существенно комплексные корни многочлена $Q(x)$ кратности β_j, $j = 1, 2, \ldots, s$, то существуют действительные числа $A_i^{(\alpha)}$, $i = 1, 2, \ldots, r$, $\alpha = 1, 2, \ldots, \alpha_i$, $M_j^{(\beta)}$ и N_j, $j = 1, 2, \ldots, s$, $\beta = 1, 2, \ldots, \beta_j$, такие, что

$$\frac{P(x)}{Q(x)} = \frac{A_1^{(1)}}{(x - a_1)^{\alpha_1}} + \frac{A_1^{(2)}}{(x - a_1)^{\alpha_1 - 1}} + \ldots + \frac{A_1^{(\alpha_1)}}{x - a_1} + \ldots$$

$$\ldots + \frac{A_r^{(1)}}{(x - a_r)^{\alpha_r}} + \frac{A_r^{(2)}}{(x - a_r)^{\alpha_r - 1}} + \ldots + \frac{A_r^{(\alpha_r)}}{x - a_r} + \frac{M_1^{(1)}x + N_1^{(1)}}{(x^2 + p_1 x + q_1)^{\beta_1}} +$$

$$+ \frac{M_1^{(2)}x + N_1^{(2)}}{(x^2 + p_1 x + q_1)^{\beta_1 - 1}} + \ldots + \frac{M_1^{(\beta_1)}x + N_1^{(\beta_1)}}{x^2 + p_1 x + q_1} + \ldots$$

$$\ldots + \frac{M_s^{(1)}x + N_s^{(1)}}{(x^2 + p_s x + q_s)^{\beta_s}} + \frac{M_s^{(2)}x + N_s^{(2)}}{(x^2 + p_s x + q_s)^{\beta_s - 1}} + \ldots + \frac{M_s^{(\beta_s)}x + N_s^{(\beta_s)}}{x^2 + p_s x + q_s}. \tag{19.30}$$

---

[1] Без ограничения общности можно считать, что коэффициент старшего члена многочлена $Q(x)$ равен единице, так как в случае, когда он равен какому-то другому числу (отличному от нуля), можно разделить числитель и знаменатель дроби $\frac{P(x)}{Q(x)}$ на это число, после чего у получившегося в знаменателе многочлена коэффициент старшего члена окажется равным единице.

Доказательство. Из разложения (19.29) имеем

$$Q(x) = (x - a_1)^{\alpha_1} Q_1(x).$$

Здесь

$$Q_1(x) = (x - a_2)^{\alpha_2} \dots (x - a_r)^{\alpha_r} (x^2 + p_1 x + q_1)^{\beta_1} \dots (x^2 + p_s x + q_s)^{\beta_s}$$

и, следовательно, $Q_1(a_1) \neq 0$, поэтому, согласно лемме 1,

$$\frac{P(x)}{Q(x)} = \frac{A_1^{(1)}}{(x - a_1)^{\alpha_1}} + \frac{P_1(x)}{(x - a_1)^{\alpha_1 - 1} Q_1(x)}.$$

Применяя в случае $\alpha_1 > 1$ подобным образом эту же лемму к рациональной дроби $\dfrac{P_1(x)}{(x - a_1)^{\alpha_1 - 1} Q_1(x)}$, получаем

$$\frac{P(x)}{Q(x)} = \frac{A_1^{(1)}}{(x - a_1)^{\alpha_1}} + \frac{A_1^{(2)}}{(x - a_1)^{\alpha_1 - 1}} + \frac{P_2(x)}{(x - a_1)^{\alpha_1 - 2} Q_2(x)}.$$

Продолжая этот процесс далее, пока показатель степени у сомножителя $x - a_1$ в знаменателе последней дроби в правой части равенства не станет равным нулю, а затем, поступая аналогичным образом относительно множителей $x - a_i$, $i = 2, \dots, r$, будем иметь

$$\frac{P(x)}{Q(x)} = \frac{A_1^{(1)}}{(x - a_1)^{\alpha_1}} + \frac{A_1^{(2)}}{(x - a)^{\alpha_1 - 1}} + \dots + \frac{A_1^{(\alpha_1)}}{x - a_1} + \dots$$

$$\dots + \frac{A_r^{(1)}}{(x - a_r)^{\alpha_r}} + \frac{A_r^{(2)}}{(x - a_r)^{\alpha_r - 1}} + \dots + \frac{A_r^{(\alpha_r)}}{x - a_r} + \frac{P^*(x)}{Q^*(x)},$$

где $\dfrac{P^*(x)}{Q^*(x)}$ — снова правильная рациональная дробь, причем $P^*(x)$ и $Q^*(x)$ — многочлены с действительными коэффициентами и многочлен $Q^*(x)$ не имеет действительных корней.

Применяя последовательно лемму 2 к дроби $\dfrac{P^*(x)}{Q^*(x)}$ и к получающимся при этом выражениям, в результате получим формулу (19.30). □

Рациональные дроби вида

$$\frac{A}{(x - a)^{\alpha}}, \quad A \neq 0, \quad \text{и} \quad \frac{Mx + N}{(x^2 + px + q)^{\beta}}, \quad M^2 + N^2 > 0,$$

где a, p, q, A, M и N — действительные числа и $p^2/4 - q < 0$ (корни трехчлена $x^2 + px + q$ существенно комплексные), называются *элементарными рациональными дробями*.

Таким образом, доказанная теорема утверждает, что *всякая ненулевая правильная рациональная дробь может быть разложена на сумму элементарных рациональных дробей*.

При выполнении разложения вида (19.30) для конкретно заданной дроби обычно оказывается удобным метод неопределенных коэффициентов. Он состоит в следующем. Для данной правильной дроби $\frac{P(x)}{Q(x)}$ записывают разложение (19.30), в котором коэффициенты $A_i^{(\alpha)}$, $M_j^{(\beta)}$, $N_j^{(\beta)}$ считают неизвестными ($i = 1, 2, \ldots, r$, $\alpha = 1, 2, \ldots, \alpha_i$, $j = 1, 2, \ldots, s$, $\beta = 1, 2, \ldots, \beta_j$). После этого обе части равенства приводят к общему знаменателю и у получившихся в числителе многочленов приравнивают коэффициенты. При этом если степень многочлена $Q(x)$ равна n, то, вообще говоря, в числителе правой части равенства (19.30) после приведения к общему знаменателю получается многочлен степени $n - 1$, т. е. многочлен с n коэффициентами, число же неизвестных $A_i^{(\alpha)}$, $M_j^{(\beta)}$, $N_j^{(\beta)}$ также равно n (см. (19.10)):

$$\sum_{i=0}^{r}\alpha_i + 2\sum_{j=0}^{s}\beta_j = n.$$

Таким образом, получена система n уравнений с n неизвестными. Существование у нее решения вытекает из доказанной теоремы.

Отметим, что после приведения выражения (19.30) к общему знаменателю и его отбрасывания, в случае когда $Q(x)$ имеет действительные корни, целесообразно подставить в обе части получившегося равенства последовательно эти корни; в результате получаются некоторые соотношения между искомыми коэффициентами, полезные для их окончательного определения.

Примеры. 1. Разложим дробь $\frac{x}{(x^2-1)(x-2)}$ на элементарные дроби. Согласно (19.30), искомое разложение имеет вид

$$\frac{x}{(x^2-1)(x-2)} = \frac{A}{x-1} + \frac{B}{x+1} + \frac{C}{x-2}.$$

Приводя дроби к общему знаменателю и отбрасывая его, получим

$$x = A(x+1)(x-2) + B(x-1)(x-2) + C(x-1)(x+1). \quad (19.31)$$

Здесь имеет место случай, когда все корни знаменателя действительны. Полагая в равенстве (19.31), согласно сказанному выше, последовательно $x = 1$, $x = -1$ и $x = 2$, находим $1 = -2A$, $-1 = 6B$, $2 = 3C$, откуда $A = -1/2$, $B = -1/6$, $C = 2/3$.

Таким образом, искомое разложение имеет вид

$$\frac{x}{(x^2-1)(x-2)} = -\frac{1}{2(x-1)} - \frac{1}{6(x+1)} + \frac{2}{3(x-2)}. \quad (19.32)$$

2. Найдем разложение на элементарные дроби для $\frac{x^2-1}{x(x^2+1)^2}$. Общий вид разложения в этом случае таков:

$$\frac{x^2-1}{x(x^2+1)^2} = \frac{A}{x} + \frac{Bx+C}{(x^2+1)^2} + \frac{Dx+E}{x^2+1}.$$

Приводя дроби к общему знаменателю и отбрасывая его, имеем

$$x^2 - 1 = A(x^2+1)^2 + (Bx+C)x + (Dx+E)(x^2+1)x.$$

Приравняем коэффициенты при одинаковых степенях x:

$$-1 = A,\ \ 0 = C + E,\ \ 1 = 2A + B + D,\ \ 0 = E,\ 0 = A + D,$$

отсюда находим: $A = -1$, $B = 2$, $C = 0$, $D = 1$, $E = 0$, поэтому искомое разложение имеет вид

$$\frac{x^2-1}{x(x^2+1)^2} = -\frac{1}{x} + \frac{2x}{(x^2+1)^2} + \frac{x}{x^2+1}. \quad (19.33)$$

Следует заметить, что в отдельных случаях разложение на элементарные дроби можно получить быстрее и проще, не прибегая к методу неопределенных коэффициентов, а каким-либо другим путем. Например, для разложения той же дроби $\frac{x^2-1}{x^2(1+x^2)^2}$ на сумму элементарных дробей проще всего дважды прибавить и вычесть в числителе x^2 и произвести деление так, как это указано ниже:

$$\frac{x^2-1}{x(x^2+1)^2} = \frac{x^2-(x^2+1)+x^2}{x(x^2+1)^2} = \frac{2x}{(x^2+1)^2} - \frac{1}{x(x^2+1)} =$$

$$= \frac{2x}{(x^2+1)^2} - \frac{(x^2+1)-x^2}{x(x^2+1)} = -\frac{1}{x} + \frac{2x}{(x^2+1)^2} + \frac{x}{x^2+1}.$$

Полученное в результате разложение и является разложением (19.33) данной дроби на сумму элементарных дробей.

УПРАЖНЕНИЕ 5. Докажите, что разложение вида (19.30) правильной рациональной дроби единственно.

З а м е ч а н и е 2. Если $\frac{P(x)}{Q(x)}$ — правильная рациональная дробь с комплексными коэффициентами, то, в силу замечания 1, применяя к ней последовательно разложение (19.25), как это было сделано при доказательстве теоремы 1 (см. замечание 1), получим разложение заданной дроби только на дроби вида $\frac{A}{(z-a)^{\alpha}}$, где a и A, вообще говоря, комплексные числа

$$\frac{P(z)}{Q(z)} = \sum_{i=1}^{k}\sum_{j=1}^{\alpha_i}\frac{A_i^{(j)}}{(z-a_i)^{\alpha_i - j + 1}}.$$

Здесь $a_1, \ldots, a_k$ — все различные корни многочлена $Q(z)$, α_i — кратность корня a_i, $i = 1, 2, \ldots, k$, поэтому сумма $\alpha_1 + \alpha_2 + \ldots \alpha_k$ равна степени многочлена $Q(z)$.

§ 20.

Интегрирование рациональных дробей

20.1. Интегрирование элементарных рациональных дробей

В этом и следующем параграфах рассмотрены методы интегрирования некоторых классов элементарных функций. При этом каждый раз, не оговаривая специально, будем предполагать, что речь идет о вычислении интеграла на некотором промежутке, во всех точках которого определена подынтегральная элементарная функция (иначе говоря, на котором формула, задающая подынтегральную функцию, имеет смысл (см. п. 5.3)).

В предыдущем параграфе показано, что всякая рациональная дробь представима в виде суммы многочлена и эле-

ментарных рациональных дробей (см. (19.24) и (19.30)). Интеграл от многочлена вычисляется, и притом очень просто (см. п. 18.2). Рассмотрим вопрос об интегрировании элементарных рациональных дробей.

Сначала рассмотрим вычисление интегралов от дробей вида $\frac{A}{(x-a)^n}$, $n = 1, 2, \ldots$.

Если $n = 1$, то (см. формулу 2 в п. 18.2)

$$\int \frac{A}{x-a}\,dx = A\ln|x-a| + C, \tag{20.1}$$

а если $n \neq 1$, то (см. формулу 1 в п. 18.2)

$$\int \frac{A}{(x-a)^n}\,dx = -\frac{A}{(n-1)(x-a)^{n-1}} + C. \tag{20.2}$$

Рассмотрим теперь интегралы от дробей $\frac{Mx+N}{(x^2+px+q)^n}$, где $\frac{p^2}{4} - q < 0$, $n = 1, 2, \ldots$. Снова начнем со случая $n = 1$. Замечая, что

$$x^2 + px + q = \left(x + \frac{p}{2}\right)^2 + \left(q - \frac{p^2}{4}\right),$$

и полагая $t = x + \frac{p}{2}$, $a^2 = q - \frac{p^2}{4} > 0$, имеем $a > 0$,

$$\int \frac{Mx+N}{x^2+px+q}\,dx = \int \frac{M\left(t - \frac{p}{2}\right) + N}{t^2+a^2}\,dt =$$

$$= M\int \frac{tdt}{t^2+a^2} + \left(N - \frac{Mp}{2}\right)\int \frac{dt}{t^2+a^2} =$$

$$= \frac{M}{2}\ln(t^2+a^2) + \frac{2N-Mp}{2a}\operatorname{arctg}\frac{t}{a} + C =$$

$$= \frac{M}{2}\ln(x^2+px+q) + \frac{2N-Mp}{2a}\operatorname{arctg}\frac{2x+p}{2a} + C. \tag{20.3}$$

В случае $n > 1$, полагая, как и выше, $t = x + \frac{p}{2}$, $a^2 = q - \frac{p^2}{4}$, подобным же образом получим

$$\int \frac{Mx+N}{(x^2+px+q)^n}\,dx = M\int \frac{tdt}{(t^2+a^2)^n} + \frac{2N-pM}{2}\int \frac{dt}{(t^2+a^2)^n}. \tag{20.4}$$

Рассмотрим отдельно каждый из получившихся интегралов в правой части этого равенства. Что касается первого из них, то он вычисляется сразу:

$$\int \frac{t\,dt}{(t^2+a^2)^n} = \frac{1}{2}\int \frac{d(t^2+a^2)}{(t^2+a^2)^n} = -\frac{1}{2(n-1)(t^2+a^2)^{n-1}} + C. \tag{20.5}$$

Второй же интеграл правой части равенства (20.4) вычисляется несколько сложнее. Пусть

$$I_n = \int \frac{dt}{(t^2+a^2)^n}, \quad n = 1, 2, 3, \ldots .$$

Проинтегрируем интеграл I_n по частям, положив $u = \frac{1}{(t^2+a^2)^n}$, $dv = dt$, и, следовательно, $du = -\frac{2nt\,dt}{(t^2+a^2)^{n+1}}$, $v = t$, а затем, добавив и вычтя a^2 в числителе получившейся под знаком интеграла функции и произведя деление так, как это указано ниже, получим

$$I_n = \int \frac{dt}{(t^2+a^2)^n} = \frac{t}{(t^2+a^2)^n} + 2n\int \frac{t^2}{(t^2+a^2)^{n+1}}\,dt =$$

$$= \frac{t}{(t^2+a^2)^n} + 2n\int \frac{(t^2+a^2)-a^2}{(t^2+a^2)^{n+1}}\,dt =$$

$$= \frac{t}{(t^2+a^2)^n} + 2n\left[\int \frac{dt}{(t^2+a^2)^n} - a^2\int \frac{dt}{(t^2+a^2)^{n+1}}\right],$$

т. е.

$$I_n = \frac{t}{(t^2+a^2)^n} + 2nI_n - 2na^2 I_{n+1},$$

откуда

$$I_{n+1} = \frac{1}{2na^2}\frac{t}{(t^2+a^2)^n} + \frac{2n-1}{2na^2} I_n, \quad n = 1, 2, \ldots . \tag{20.6}$$

Интеграл I_1 легко вычисляется (см. в п. 18.3 формулу 12); формула (20.6) позволяет вычислить I_2; зная же I_2, по той же формуле можно найти значение I_3, продолжая этот процесс дальше, можно найти выражение для любого интеграла I_n $(n = 1, 2, \ldots)$.

20.2. Общий случай

Из результатов п. 19.6 и предыдущего п. 20.1 непосредственно вытекает следующая теорема.

ТЕОРЕМА 1. *Неопределенный интеграл от любой рациональной дроби на всяком промежутке, на котором знаменатель дроби не обращается в нуль, существует и выражается через элементарные функции, а именно он является алгебраической суммой суперпозиций рациональных дробей, арктангенсов и натуральных логарифмов.*

Теорема 1 есть прямое следствие формул (19.24), (19.30), (18.7), (18.8), (18.9), (20.1)—(20.6). Эти формулы дают и конкретный способ вычисления интеграла от рациональной функции: сначала делением числителя на знаменатель выделяется «целая часть», т. е. данная рациональная дробь представляется в виде суммы многочлена и правильной рациональной дроби (19.24). Если получившаяся правильная рациональная дробь оказывается ненулевой, то она раскладывается на сумму элементарных дробей (19.30), после чего, используя линейность интеграла (18.7) и (18.8), можно вычислить интегралы от каждого слагаемого отдельно, согласно формулам (18.9) и (20.1)—(20.6).

Примеры. 1. Вычислим $\int \frac{x\,dx}{(x^2-1)(x-2)}$. Уже известно (см. (19.32)), что

$$\frac{x}{(x^2-1)(x-2)} = -\frac{1}{2(x-1)} - \frac{1}{6(x+1)} + \frac{2}{3(x-2)},$$

поэтому

$$\int \frac{x\,dx}{(x^2-1)(x-2)} = -\frac{1}{2}\int \frac{dx}{x-1} - \frac{1}{6}\int \frac{dx}{x+1} + \frac{2}{3}\int \frac{dx}{x-2} =$$

$$= -\frac{1}{2}\ln|x-1| - \frac{1}{6}\ln|x+1| + \frac{2}{3}\ln|x-2| + C.$$

2. Вычислим $\int \frac{x^6+2x^4+2x^2-1}{x(x^2+1)^2}\,dx$. Согласно общему правилу, выделим целую часть, разделив числитель на знаменатель; имеем

$$\frac{x^6+2x^4+2x^2-1}{x(x^2+1)^2} = x + \frac{x^2-1}{x(x^2+1)^2}.$$

Для получившейся правильной рациональной дроби уже найдено ее разложение на элементарные дроби (см. формулу (19.33)):

$$\frac{x^2-1}{x(x^2+1)^2} = -\frac{1}{x} + \frac{2x}{(x^2+1)^2} + \frac{x}{x^2+1},$$

поэтому

$$\int \frac{x^6+2x^4+2x^2-1}{x(x^2+1)^2}\,dx = \int x dx - \int \frac{dx}{x} + \int \frac{2xdx}{(x^2+1)^2} + \int \frac{x}{x^2+1}\,dx =$$

$$= \frac{x^2}{2} - \ln|x| + \int \frac{d(x^2+1)}{(x^2+1)^2} + \frac{1}{2}\int \frac{d(x^2+1)}{x^2+1} =$$

$$= \frac{x^2}{2} - \ln|x| - \frac{1}{x^2+1} + \frac{1}{2}\ln(x^2+1) + C.$$

Следует иметь в виду, что указанный метод вычисления неопределенного интеграла от рациональной дроби является общим: с помощью его можно вычислить неопределенный интеграл от любой рациональной дроби, если можно получить конкретное разложение знаменателя на множители вида (19.10). Однако естественно, что в отдельных частных случаях бывает целесообразнее для значительного сокращения вычислений действовать иными путями.

Например, для вычисления интеграла $I = \int \frac{x^2dx}{(1-x^2)^3}$ проще не раскладывать подынтегральную функцию на элементарные дроби, а применить правило интегрирования по частям. Положив $u = x$, $dv = \frac{xdx}{(1-x^2)^3}$ и, следовательно, $du = dx$, $v = \frac{1}{4(1-x^2)^2}$, получим

$$I = -\frac{1}{2}\int x\frac{d(1-x^2)}{(1-x^2)^3} = \frac{x}{4(1-x^2)^2} - \frac{1}{4}\int \frac{1}{(1-x^2)^2}\,dx.$$

Прибавляя и вычитая в числителе получившейся подынтегральной функции x^2, производя деление, получаем два интеграла, из которых первый табличный, а второй легко вычисляется интегрированием по частям:

$$I = \frac{x}{4(1-x^2)^2} - \frac{1}{4}\int\frac{(1-x^2)+x^2}{(1-x^2)^2}dx = \frac{x}{4(1-x^2)^2} - \frac{1}{4}\int\frac{dx}{1-x^2} - \frac{1}{4}\int\frac{x^2dx}{(1-x^2)^2} =$$

$$= \frac{x}{4(1-x^2)^2} - \frac{1}{8}\ln\left|\frac{1+x}{1-x}\right| + \frac{1}{8}\int x\frac{d(1-x^2)}{(1-x^2)^2} =$$

$$= \frac{x}{4(1-x^2)^2} - \frac{1}{8}\ln\left|\frac{1+x}{1-x}\right| - \frac{x}{8(1-x^2)} + \frac{1}{8}\int\frac{dx}{1-x^2} =$$

$$= \frac{x}{4(1-x^2)^2} - \frac{1}{16}\ln\left|\frac{1+x}{1-x}\right| - \frac{x}{8(1-x^2)} + C.$$

20.3*. Метод Остроградского

В п. 19.6 было показано, что всякая правильная ненулевая рациональная дробь может быть представлена в виде суммы элементарных дробей. Но из п. 20.1 следует, что первообразные элементарных дробей $\frac{1}{x-a}$ и $\frac{Mx+N}{x^2+px+q}\left(\frac{p^2}{4}-q<0\right)$ являются трансцендентными функциями вида $A\operatorname{arctg}(a_1x+a_2) + B\ln(b_1x+b_2) + C$ (см. 20.1) и (20.3)); первообразная элементарной дроби $\frac{1}{(x-a)^\alpha}$, $\alpha = 2, 3, \dots$, является рациональной дробью; первообразная же элементарной дроби $\frac{(Mx+N)}{(x^2+px+q)^\beta}$, $\beta = 2, 3, \dots$, в силу формул (20.4), (20.5), (20.6) и формулы 12 п. 18.3 может быть, вообще говоря, представлена в виде суммы правильной рациональной дроби и трансцендентной функции вида $A\operatorname{arctg}(a_1x+a_2)+C$, являющейся первообразной от дроби вида $\frac{B}{x^2+px+q}\left(\frac{p^2}{4}-q<0\right)$. Поэтому всякая первообразная любой рациональной дроби представима, вообще говоря, в виде суммы рациональной дроби (алгебраическая часть) и трансцендентной функции, являющейся первообразной от суммы дробей вида $\frac{A}{x-a}$ и $\frac{Mx+N}{x^2+px+q}$, $\frac{p^2}{4}-q<0$.

Таким образом, если $\frac{P(x)}{Q(x)}$ — правильная рациональная дробь и

$$Q(x) = (x - a_1)^{\alpha_1} \dots (x - a_r)^{a_r}(x^2 + p_1 x + q_1)^{\beta_1} \dots (x^2 + p_s x + q_s)^{\beta_s}$$

— разложение ее знаменателя в виде (19.10), то существуют такие многочлены $P_1(x)$, $Q_1(x)$ и такие числа A_i, $i = 1, 2, \dots, r$, M_j, N_j, $j = 1, 2, \dots s$, что имеет место равенство

$$\int \frac{P(x)}{Q(x)}\, dx =$$

$$= \frac{P_1(x)}{Q_1(x)} + \int \left[\sum_{i=1}^{r} \frac{A_i}{x - a_i} + \sum_{i=1}^{s} \frac{M_j x + N_j}{x^2 + p_j x + q_j} \right] dx; \qquad (20.7)$$

отсюда, произведя под знаком интеграла сложение дробей, имеем

$$\int \frac{P(x)}{Q(x)}\, dx = \frac{P_1(x)}{Q_1(x)} + \int \frac{P_2(x)}{Q_2(x)}\, dx, \qquad (20.8)$$

где $Q_2(x) = (x - a_1) \dots (x - a_r)(x^2 + p_1 x + q_1) \dots (x_2 + p_s x + q_s)$. Из формул (20.2) и (20.6) следует, что многочлен $Q_1(x)$ имеет вид

$$Q_1(x) =$$

$$= (x - a_1)^{\alpha_1 - 1} \dots (x - a_r)^{\alpha_r - 1}(x^2 + p_1 x + q_1)^{\beta_1 - 1} \dots (x^2 + p_s x + q_s)^{\beta_s - 1},$$

т. е. является наибольшим общим делителем многочлена $Q(x)$ и его производной $Q'(x)$ (см. (19.23)).

Формула (20.8) называется формулой *Остроградского*[1]. Второе слагаемое правой части формулы (20.8) называется трансцендентной частью интеграла $\int \frac{P(x)}{Q(x)}\, dx$; это естественно, ибо из сказанного выше следует, что всякая первообразная дроби $P_2(x)/Q_2(x)$ с точностью до постоянного слагаемого представляет собой линейную комбинацию логарифмов и арктангенсов от рациональных функций и, значит, как это можно показать, будет являться, вообще говоря, трансцендентной функцией. Первое же слагаемое, называемое алгебраической частью, может быть найдено чисто алгебраическим путем, если известны многочлены $P(x)$ и $Q(x)$ (а значит, и $Q'(x)$), т. е. без интегрирования каких-либо функций.

[1] М. В. Остроградский (1801—1869) — русский математик.

В самом деле, многочлен $Q_1(x)$, являясь наибольшим общим делителем многочленов $Q(x)$ и $Q'(x)$, всегда может быть найден с помощью алгоритма Евклида (см. п. 19.5*), тем самым для отыскания многочлена $Q_1(x)$ не требуется знания корней многочлена $Q(x)$; однако, если корни многочлена $Q(x)$ известны, а значит, известно и его разложение вида (19.17), то ясно, что многочлен $Q_1(x)$ может быть сразу записан по формуле (19.23). Многочлен $Q_2(x)$ находится как частное от деления $Q(x)$ на $Q_1(x)$.

Для отыскания же многочленов $P_1(x)$ и $P_2(x)$ можно применить метод неопределенных коэффициентов. Поясним его. Обозначим степень многочлена $Q_1(x)$ через n_1 — степень многочлена $Q_2(x)$ — через n_2; тогда из равенства

$$Q(x) = Q_1(x)Q_2(x) \tag{20.9}$$

получим $n = n_1 + n_2$. В силу того, что дроби $\frac{P_1(x)}{Q_1(x)}$ и $\frac{P_2(x)}{Q_2(x)}$, являясь суммой правильных рациональных дробей, также правильные, степени многочленов $P_1(x)$ и $P_2(x)$ соответственно не выше, чем $n_1 - 1$ и $n_2 - 1$ и, значит, в этих многочленах число отличных от нуля коэффициентов соответственно не превышает n_1 и n_2; таким образом, число неизвестных коэффициентов равно $n_1 + n_2 = n$. Дифференцируя первообразные, входящие в обе части формулы (20.8), получим (опуская для краткости обозначение аргумента) соотношение

$$\frac{P}{Q} = \left(\frac{P_1}{Q_1}\right)' + \frac{P_2}{Q_2}.$$

Дифференцируя дробь $\frac{P_1}{Q_1}$, имеем

$$\frac{P}{Q} = \frac{P_1'Q_1 - P_1Q_1'}{Q_1^2} + \frac{P_2}{Q_2}. \tag{20.10}$$

Заметим, что

$$\frac{P_1'Q_1' - P_1Q_1'}{Q_1^2} = \frac{P_1'Q_2 - P_1R}{Q_1Q_2}, \tag{20.11}$$

где $R = \frac{Q'_1Q_2}{Q_1}$ — многочлен. Действительно, если z — корень многочлена Q_1 кратности λ, то, как известно (см. п. 19.4),

z является корнем кратности $\lambda - 1$ для производной Q'_1 и однократным корнем многочлена Q_2, поэтому в этом случае z является и корнем кратности λ для многочлена $Q'_1 Q_2$. Отсюда, согласно формуле (19.7), сразу следует, что многочлен $Q'_1 Q_2$ нацело делится на многочлен Q_1, т. е. что R также многочлен. Итак, из формул (20.9), (20.10) и (20.11) имеем

$$\frac{P}{Q} = \frac{P'_1 Q_2 - P_1 R}{Q_1 Q_2} + \frac{P_2}{Q_2},$$

откуда

$$P = P'_1 Q_2 - P_1 R + P_2 Q_1. \tag{20.12}$$

Многочлен P имеет степень не выше чем $n - 1$ $\left(\text{так как дробь } \frac{P}{Q} \text{ — правильная}\right)$. Приравняв коэффициенты при одинаковых степенях k, $k = 0, 1, \ldots, n - 1$, переменного x в обеих частях равенства (20.12), получим n линейных уравнений относительно n неизвестных. Выше было доказано (см. (20.8)), что многочлены P_1 и P_2 всегда (в частности, при некотором фиксированном многочлене Q и при любом многочлене P степени, не превышающей $n - 1$) существуют; поэтому полученная система линейных уравнений имеет решение при любой правой части[1]. Отсюда следует, что определитель этой системы не равен нулю, а значит, про рассматриваемую систему можно сказать, что она не только имеет решение, но и что оно единственно. Тем самым не только получен метод определения неизвестных коэффициентов в формуле (20.8), но и доказана единственность этого представления.

Формула (20.8) сводит, вообще говоря, задачу интегрирования любой правильной рациональной дроби к задаче интегрирования правильной рациональной дроби, у которой знаменатель имеет только простые корни. С помощью этой формулы при интегрировании правильной рациональной дроби можно найти указанным выше способом алгебраическую часть интеграла $\int \frac{P(x)}{Q(x)}\,dx$, а затем проинтегрировать бо-

---

[1] Как обычно, предполагается, что все члены уравнений, содержащие неизвестные, и только они перенесены в левую часть равенства.

лее простую рациональную дробь $\frac{P_2(x)}{Q_2(x)}$, если, конечно, случайно не окажется, что $P_2(x)$ — тождественный нуль; в этом случае задача будет уже решена.

Описанный метод интегрирования рациональных дробей называется *методом Остроградского*. При использовании метода Остроградского для интегрирования рациональных дробей часто оказывается целесообразней записывать формулу Остроградского (20.8) в виде (20.7), так как в этом случае после нахождения неизвестных коэффициентов в подынтегральной функции ее сразу можно проинтегрировать.

Неизвестные коэффициенты в формуле (20.7) находят методом, который описан для формулы (20.8): следует продифференцировать обе части равенства (20.7), привести к общему знаменателю все рациональные дроби, получившиеся в обеих частях равенства, приравнять коэффициенты у одинаковых степеней переменной x в многочленах, стоящих в числителях, и решить получившуюся систему линейных уравнений.

Пример. Применим метод Остроградского для вычисления интеграла $\int \frac{x^4 + 2x^3 - 2x^2 + x}{(1 - x)^3(1 + x^2)^2}\,dx$. Здесь

$$Q(x) = (1 - x)^3(1 + x^2)^2.$$

Общий наибольший делитель $Q_1(x)$ многочлена $Q(x)$ и его производной $Q'(x)$, согласно формуле (19.23), имеет вид

$$Q_1(x) = (1 - x)^2(1 + x^2) = x^4 - 2x^3 + 2x^2 - 2x + 1.$$

Поэтому

$$Q_2(x) = \frac{Q(x)}{Q_1(x)} = (1 - x)(1 + x^2)$$

(что, конечно, соответствует формуле для $Q_2(x)$ в равенстве (20.8)). Нам еще понадобится

$$Q_1'(x) = -2(1 - x)(1 + x^2) + 2(1 - x)^2x = -2(1 - x)(2x^2 - x + 1).$$

Многочлены $P_1(x)$ и $P_2(x)$ запишем с неопределенными коэффициентами так, чтобы дроби $\frac{P_1(x)}{Q_1(x)}$ и $\frac{P_2(x)}{Q_2(x)}$ были правильными:

$$P_1(x) = Kx^3 + Lx^2 + Mx + N,\ P_2(x) = kx^2 + lx + m.$$

Согласно формуле (20.8),

$$\int \frac{x^4 + 2x^3 - 2x^2 + x}{(1-x)^3(1+x^2)^2}\,dx =$$

$$= \frac{Kx^3 + Lx^2 + Mx + N}{(1-x)^2(1+x^2)} + \int \frac{kx^2 + lx + m}{(1-x)(1+x^2)}\,dx,$$

поэтому

$$\frac{x^4 + 2x^3 - 2x^2 + x}{(1-x)^3(1+x^2)^2} = \left[\frac{Kx^2 + Lx^2 + Mx + N}{(1-x)^2(1+x^2)}\right]' + \frac{kx^2 + lx + m}{(1-x)(1+x^2)}.$$

Произведя дифференцирование и используя полученную формулу для $Q'(x)$, будем иметь

$$\frac{x^4 + 2x^3 - 2x^2 + x}{(1-x)^3(1+x^2)^2} =$$

$$= \frac{(3Kx^2 + 2Lx + M)(1-x)^2(1+x^2)}{(1-x)^4(1-x^2)^2} +$$

$$+ \frac{2(Kx^3 + Lx^2 + Mx + N)(1-x)(2x^2 - x + 1)}{(1-x)^4(1-x^2)^2} + \frac{kx^2 + lx + m}{(1-x)(1+x^2)}.$$

Сократив первую дробь в правой части равенства на множитель $1 - x$ и приведя все дроби к общему знаменателю, приравняем числители левой и правой части равенства:

$$x^4 + 2x^3 - 2x^2 + x = (3Kx^2 + 2Lx + M)(-x^3 + x^2 - x + 1) +$$

$$+ 2(Kx^3 + Lx^2 + Mx + N)(2x^2 - x + 1) +$$

$$+ (kx^2 + lx + m)(x^4 - 2x^3 + 2x^2 - 2x + 1).$$

Отметим, что это равенство следует и из формулы (20.12).

Приравняв коэффициенты при одинаковых степенях переменной x в обеих частях равенства, получаем следующую систему уравнений:

$$M + 2N + m = 0,$$

$$2L - M + 2M - 2N - 2m + l = 1,$$

$$3K - 2L + M + 2L - 2M + 4N + k - 2l + 2m = -2,$$

$$-3K + 2L - M + 2K - 2L + 4M - 2k + 2l - 2m = 2,$$

$$3K - 2L - 2K + 4L + 2k - 2l + m = 1,$$

$$-3K + 4K - 2k + l = 0, \quad k = 0,$$

или

$$M + 2N + m = 0,$$
$$2L + M - 2N + l - 2m = 1,$$
$$3K - M + 4N + k - 2l + 2m = -2,$$
$$-K + 3M - 2k + 2l - 2m = 2,$$
$$K + 2L + 2k - 2l + m = 1,$$
$$K - 2k + l = 0,\ \ k = 0.$$

Решая эту систему уравнений, находим: $K = 1/2$, $L = -1/2$, $M = 3/2$, $N = -1$, $k = 0$, $l = -1/2$, $m = 1/2$; поэтому

$$\int \frac{x^4 + 2x^3 - 2x^2 + x}{(1-x)^3(x^2+1)^2}\,dx =$$
$$= \frac{1}{2}\,\frac{x^3 - x^2 + 3x - 2}{2(1-x)^2(1+x^2)} + \frac{1}{2}\int \frac{-x+1}{(1-x)(1+x^2)}\,dx =$$
$$= \frac{x^3 - x^2 + 3x - 2}{2(1-x)^2(1+x^2)} + \frac{1}{2}\operatorname{arctg} x + C.$$

§ 21.

Интегрирование некоторых иррациональностей

21.1. Предварительные замечания

Функции вида

$$R(u_1, \dots, u_n) = \frac{P(u_1, \dots, u_n)}{Q(u_1, \dots, u_n)}, \tag{21.1}$$

где P и Q — многочлены от переменных $u_1, \dots, u_n$, т. е. функции вида

$$\sum_{k_1 + \dots + k_n \leqslant k}^{n-1} a_{k_1, \dots, k_n} u_1^{k_1} \dots u_n^{k_n},$$

называются *рациональными функциями от* $u_1, \dots, u_n$.

Если в формуле (21.1) переменные $u_1, \dots, u_n$ в свою очередь являются функциями переменной x: $u_i = \varphi_i(x)$, $i = 1, 2, \dots, n$, то функция $R[\varphi_1(x), \dots, \varphi_n(x)]$ называется *рациональной функцией* от функций $\varphi_1(x), \dots, \varphi_n(x)$.

Например, функция $f(x) = \dfrac{x + \sqrt[3]{(x^2-1)^2}}{\sqrt{x} - \sqrt{x^2+1}}$ является рациональной функцией от x и радикалов $\sqrt{x}$, $\sqrt[3]{x^2-1}$, и $\sqrt{x^2+1}$:

$$f(x) = R(x, \sqrt{x}, \sqrt[3]{x^2-1}, \sqrt{x^2+1});$$

здесь

$$R(u_1, u_2, u_3, u_4) = \frac{u_1 + u_3^2}{u_2 - u_4},$$

$$u_1 = x, \quad u_2 = \sqrt{x}, u_3 = \sqrt[3]{x^2-1}, \quad u_4 = \sqrt{x^2+1}.$$

Если в формуле (21.1) переменные $u_1, \ldots, u_n$ являются элементарными тригонометрическими функциями, то получающаяся сложная функция называется рациональной относительно элементарных тригонометрических функций. Примером такой функции является следующая:

$$\frac{\sin^2 x - \cos^2 x}{\sin^3 x + \cos x} = R(\sin x, \cos x).$$

Перейдем теперь к интегралам от функций рассмотренных типов и покажем, что в ряде случаев они сводятся к интегралам от рациональных функций.

21.2. Интегралы вида $\int R\left[x, \left(\frac{ax+b}{cx+d}\right)^{r_1}, \ldots, \left(\frac{ax+b}{cx+d}\right)^{r_s}\right] dx$

Рассмотрим интегралы, указанные в заглавии пункта, при условии, что постоянные $r_1, \ldots, r_s$ рациональны, и $\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$ (a, b, c, d — постоянные). Последнее предположение естественно, так как если $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = 0$, то коэффициенты a, b были бы пропорциональны коэффициентам c, d и поэтому отношение $\frac{ax+b}{cx+d}$ не зависело бы от x. Подынтегральная функция в этом случае была бы обыкновенной рациональной дробью от одного переменного, вопрос об интегрировании которой был рассмотрен выше.

Пусть m — общий знаменатель чисел $r_1, \dots, r_s$:

$$r_i = \frac{p_i}{m}, \quad p_i \text{ — целое}, \quad i = 1, 2, \dots, s.$$

Положим

$$t^m = \frac{ax + b}{cx + d}, \tag{21.2}$$

откуда

$$x = \frac{dt^m - b}{a - ct^m} = \rho(t); \tag{21.3}$$

$\rho(t)$ является рациональной функцией, поэтому $\rho'(t)$ также рациональная функция; далее,

$$dx = \rho'(t)\, dt, \tag{21.4}$$

$$\left(\frac{ax + b}{cx + d}\right)^{r_i} = t^{mr_i} = t^{p_i}, \; i = 1, 2, \dots, s. \tag{21.5}$$

Подставляя (21.3), (21.4) и (21.5) в подынтегральное выражение рассматриваемого интеграла, получим

$$\int R\left[x, \left(\frac{ax + b}{cx + d}\right)^{r_1}, \dots, \left(\frac{ax + b}{cx + d}\right)^{r_s}\right] dx =$$

$$= \int R\left(\frac{dt^m - b}{a - ct^m}, t^{p_1}, \dots, t^{p_s}\right)\rho'(t)\, dt = \int R^*(t)\, dt,$$

где $R^*(t) = R\left(\frac{dt^m - b}{a - ct^m}, t^{p_1}, \dots, t^{p_s}\right)\rho'(t)$, очевидно, является рациональной функцией переменного t. Таким образом, вычисление интеграла

$$\int R\left[x, \left(\frac{ax + b}{cx + d}\right)^{r_i}, \dots, \left(\frac{ax + b}{cx + d}\right)^{r_s}\right] dx \tag{21.6}$$

сводится к интегрированию рациональных дробей.

Конечно, для того чтобы найти выражение для исходного интеграла, надо после вычисления интеграла $\int R^*(t)\, dt$, сделав обратную замену переменного $t = \left(\frac{ax + b}{cx + d}\right)^{1/m}$, вернуться к первоначальной переменной x. В дальнейшем в аналогичных ситуациях мы не будем каждый раз оговаривать необходимость обратного перехода к исходной переменной x.

Отметим, что к рассмотренному типу интегралов относятся интегралы вида

$$\int R[x, (ax+b)^{r_1}, \dots, (ax+b)^{r_s}]\,dx, \ a \neq 0,$$

в частности

$$\int R(x, x^{r_1}, \dots, x^{r_s})\,dx.$$

Пример. Вычислим интеграл $\int \frac{dx}{\sqrt{x} + \sqrt[3]{x}}$. Полагая, согласно общему правилу, $x = t^6$, $dx = 6t^5\,dt$, получаем

$$\int \frac{dx}{\sqrt{x} + \sqrt[3]{x}} = 6\int \frac{t^3}{t+1}\,dt = 6\left[\int (t^2 - t + 1)\,dt - \int \frac{dt}{t+1}\right] =$$

$$= 6\left[\frac{t^3}{3} - \frac{t^2}{2} + t - \ln|t+1|\right] + C =$$

$$= 2\sqrt{x} - 3\sqrt[3]{x} + 6\sqrt[6]{x} - 6\ln(\sqrt[6]{x} + 1) + C.$$

К интегралам вида (21.6) сводятся иногда с помощью элементарных преобразований и интегралы других типов, например, типа $\int \sqrt{(x-a)(x-b)}\,dx$.

Покажем метод вычисления подобных интегралов на примере интеграла

$$\int \sqrt{(x-1)(x-2)}\,dx. \tag{21.7}$$

Вынося в подынтегральной функции множитель $(x - 1)$ за знак радикала, получаем интеграл вида (21.6): именно при $x \geqslant 2$

$$\int \sqrt{(x-1)(x-2)}\,dx = \int (x-1)\sqrt{\frac{x-2}{x-1}}\,dx,$$

а при $x < 1$

$$\int \sqrt{(x-1)(x-2)}\,dx = \int (1-x)\sqrt{\frac{x-2}{x-1}}\,dx.$$

При $1 < x < 2$ подынтегральное выражение чисто мнимое.

Рассмотрим, например, случай $x \geqslant 2$. Положим здесь (см. (21.2)) $t^2 = \frac{x-2}{x-1}$, тогда

$$x = \frac{2-t^2}{1-t^2}, \quad dx = \frac{2t\,dt}{(1-t^2)^2},$$

поэтому

$$\int \sqrt{(x-1)(x-2)}\,dx =$$
$$= \int \left(\frac{2-t^2}{1-t^2} - 1\right)\frac{2t^2dt}{(1-t^2)^2} = 2\int \frac{t^2dt}{(1-t^2)^3}$$

— получился интеграл от рациональной дроби, который был вычислен раньше (см. п. 20.2).

21.3. Интегралы вида $\int R(x, \sqrt{ax^2+bx+c})\,dx$. Подстановки Эйлера

Указанные интегралы могут быть сведены с помощью замены переменного к интегралам от рациональных функций. Рассмотрим три замены переменного, называющиеся *подстановками Эйлера*[1]. Итак, пусть дан интеграл

$$\int R(x, \sqrt{ax^2+bx+c})\,dx,\ a \neq 0. \tag{21.8}$$

Первый случай: $a > 0$. Выполним замену x на t следующим образом:

$$\sqrt{ax^2+bx+c} = \pm x\sqrt{a} \pm t \tag{21.9}$$

(знаки можно брать в любой комбинации). Возведем обе части написанного равенства в квадрат:

$$ax^2+bx+c = ax^2 \pm tx\sqrt{a} + t^2,$$

отсюда

$$x = \frac{t^2-c}{b \mp 2t\sqrt{a}} = R_1(t),$$

$R_1(t)$ — рациональная функция от t, значит, $R_1'(t)$ также рациональная функция.

Далее, $dx = R_1'(t)\,dt$,

$$\sqrt{ax^2+bx+c} = \pm R_1(t)\sqrt{a} \pm t = R_2(t),$$

где, очевидно, $R_2(t)$ — рациональная функция. Окончательно,

$$\int R(x, \sqrt{ax^2+bx+c})\,dx =$$
$$= \int R(R_1(t),\ R_2(t))R_1'(t)\,dt = \int R^*(t)\,dt,$$

где $R^*(t) = R(R_1(t), R_2(t))\ R_1'(t)$ — рациональная дробь. □

---

[1] Л. Эйлер (1707—1783) — швейцарский математик.

В т о р о й с л у ч а й: корни трехчлена $ax^2 + bx + c$ действительные. Пусть x_1 и x_2 действительные и являются корнями трехчлена $ax^2 + bx + c$. Если $x_1 = x_2$, то

$$\sqrt{ax^2 + bx + c} = \sqrt{a(x - x_1)^2} = |x - x_1|\sqrt{a}\,.$$

Отсюда следует, что в этом случае либо под корнем стоит отрицательная при всех значениях $x \neq x_1$ величина, т. е. корень принимает только чисто мнимые выражения, — этот случай имеет место при $a < 0$ и мы его не рассматриваем, либо при $a \geqslant 0$ после указанного элементарного преобразования получаем, что переменное x не входит под знак корня, т. е. под интегралом стоит просто рациональная функция от x, вообще говоря, разная для каждого из промежутков $(-\infty, x_1)$ и $(x_1, +\infty)$.

Рассмотрим теперь случай, когда $x_1 \neq x_2$. Замечая, что $ax^2 + bx + c = a(x - x_1)(x - x_2)$, и, вынося $x - x_1$ из-под знака корня, получаем, что

$$R(x, \sqrt{ax^2 + bx + c}) = R\left(x, |x - x_1|\sqrt{\frac{a(x - x_2)}{x - x_1}}\right) =$$

$$= R_3\left(x, \sqrt{\frac{a(x - x_2)}{x - x_1}}\right), \qquad (21.10)$$

здесь $R_3(u, v)$ — рациональная функция переменных u и v.

Как известно (см. п. 21.1), интеграл от функции (21.10) может быть с помощью подстановки (см. 21.2) $t^2 = \dfrac{a(x - x_2)}{x - x_1}$ сведен к интегралу от рациональной функции; в нашем случае

$$\pm(x - x_1)t = \sqrt{a(x - x_1)(x - x_2)}\,,$$

или, беря $t > 0$ при $x \geqslant x_1$ и $t < 0$ при $x \leqslant x_1$, имеем

$$(x - x_1)t = \sqrt{ax^2 + bx + c}\,. \;\square$$

Рассмотренный в предыдущем пункте интеграл (21.7) является примером случая 2; этот интеграл был сведен выше к рациональной дроби приемом, разобранным сейчас в общем случае.

Два изученных нами способа вычисления интеграла (21.8) позволяют всегда свести этот интеграл к интегралу от раци-

ональной дроби на любом промежутке, если только корень $\sqrt{ax^2+bx+c}$ на этом промежутке не принимает чисто мнимые значения (естественно, изучая анализ в действительной области, исключить этот случай из рассмотрения). В самом деле, допустим, что ни первый, ни второй случай не имеют места, т. е. $a<0$ и корни x_1 и x_2 трехчлена ax^2+bx+c существенно комплексны: $x_1=g+hi$, $x_2=g-hi$, $h\neq 0$. Тогда

$$\sqrt{ax^2+bx+c} = \sqrt{a(x-x_1)(x-x_2)} =$$

$$= \sqrt{a(x-g-hi)(x-g+hi)} = \sqrt{a[(x-g)^2+h^2]},$$

и так как $a<0$, а $h\neq 0$, то под корнем при любых x стоит отрицательное выражение. □

Т р е т и й с л у ч а й: $c>0$. В этом случае можно применить подстановку

$$\sqrt{ax^2+bx+c} = \pm\sqrt{c} \pm xt$$

(комбинация знаков произвольна). Возводя в квадрат, получим равенство

$$ax^2+bx = \pm 2\sqrt{c}\,xt + x^2t^2,$$

откуда

$$x = \frac{b\mp 2t\sqrt{c}}{t^2-a} = R_4(t),\ dx = R_4'(t)\,dt,$$

$$\sqrt{ax^2+bx+c} = \pm\sqrt{c} \pm R_4(t)t = R_5(t),$$

где $R_4(t)$, $R_4'(t)$ и $R_5(t)$ — суть рациональные функции t. Поэтому

$$\int R(x, \sqrt{ax^2+bx+c})\,dx =$$

$$= \int R(R_4(t), R_5(t))\,R_4'(t)\,dt = \int \tilde{R}(t)\,dt,$$

где $\tilde{R}(t) = R(R_4(t), R_5(t))\,R_4'(t)$ — рациональная дробь. □

Интегралы вида $\int R(x, \sqrt{ax+b}, \sqrt{cx+d})\,dx$ сводятся подстановкой

$$t^2 = ax+b \qquad (21.11)$$

к рассмотренным интегралам вида (21.8).

В самом деле, из (21.11) имеем

$$x=\frac{t^2-b}{a},\quad dx=\frac{2}{a}t\,dt,\ \sqrt{cx+d}=\sqrt{\frac{c}{a}t^2-\frac{cb}{a}+d}=\sqrt{At^2+B},$$

где $A=\frac{c}{a}$, $B=-\frac{cb}{a}+d$, поэтому

$$\int R(x,\sqrt{ax+b},\ \sqrt{cx+d})\,dx=\int R_6(t,\sqrt{At^2+B})\,dt,$$

где $R_6(u,\ v)$ — рациональная функция переменных u и v. В правой части последнего равенства стоит интеграл типа (21.8). □

Вычисление интегралов с помощью подстановок Эйлера обычно приводит к громоздким выражениям, поэтому их следует применять, вообще говоря, лишь тогда, когда рассматриваемый интеграл не удается вычислить другим, более коротким способом. Например, замечая, что $ax^2+bx+c=$ $=a\left(x+\frac{b}{2a}\right)^2+c-\frac{b^2}{4a}$, нетрудно убедиться, что интеграл (21.8) в случае, когда подкоренное выражение положительно на некотором интервале, с помощью линейной подстановки может быть приведен (ср. с п. 18.4) к одному из трех интегралов:

$$\int R(t,\sqrt{1-t^2})\,dt,\ \int R(t,\sqrt{t^2-1})\,dt,\ \int R(t,\sqrt{t^2+1})\,dt.$$

Для вычисления полученных интегралов часто оказывается очень удобным использовать тригонометрические подстановки $t=\sin u$, $t=\cos u$, $t=\operatorname{tg} u$, а также гиперболические подстановки

$$t=\operatorname{sh} u,\ t=\operatorname{ch} u,\ t=\operatorname{th} u.$$

Пример. Вычислим интеграл $\int\sqrt{1+x^2}\,dx$ с помощью замены переменного $x=\operatorname{sh} t$, из которой следует (см. п. 9.8), что $t=\ln(x+\sqrt{x^2+1})$. Имеем

$$\int\sqrt{1+x^2}\ dx=\int\operatorname{ch}^2 t\,dt=\int\frac{1+\operatorname{ch}2t}{2}\ dt=$$

$$=\frac{t}{2}+\frac{1}{4}\operatorname{sh}2t+C=\frac{1}{2}(t+\operatorname{sh}t\operatorname{ch}t)+C=$$

$$=\frac{1}{2}[\ln(x+\sqrt{x^2+1})+x\sqrt{1+x^2}]+C.$$

21.4. Интегралы от дифференциальных биномов

Выражение $x^m(a + bx^n)^p\,dx$ ($a \neq 0$, $b \neq 0$) называется *дифференциальным биномом*. Будем рассматривать случаи, когда n, m и p — рациональные, а a и b — действительные числа.

Положим

$$x = t^{1/n}, \tag{21.12}$$

тогда $dx = \frac{1}{n} t^{1/n - 1}\,dt$ и, следовательно,

$$\int x^m(a + bx^n)^p\,dx = \frac{1}{n}\int (a + bt)^p\,t^{\frac{m+1}{n} - 1}\,dt.$$

Таким образом, интеграл

$$\int x^m(a + bx^n)^p\,dx \tag{21.13}$$

сводится подстановкой (21.12) к интегралу типа

$$\int (a + bt)^p t^q\,dt, \tag{21.14}$$

где p и q рациональны. В рассматриваемом случае

$$q = \frac{m+1}{n} - 1.$$

Первый случай: p — целое число. Пусть $q = r/s$, где r и s — целые числа. Согласно результатам п. 21.2, в этом случае подстановка $z = t^{1/s}$ сводит интеграл (21.14) к интегралу от рациональной дроби.

Второй случай: q — целое число. Пусть теперь $p = r/s$, r и s — целые числа. Согласно результатам п. 21.2, интеграл (21.14) приводится в этом случае подстановкой $z = (a + bt)^{1/s}$ к интегралу от рациональной дроби.

Третий случай: $p + q$ — целое. Пусть $p = r/s$, где s — целое. Запишем для наглядности интеграл (21.14) в виде

$$\int (a + bt)^p t^q\,dt = \int \left(\frac{a + bt}{t}\right)^p t^{p+q}\,dt.$$

Снова имеем интеграл типа, рассмотренного в том же п. 21.2. На этот раз к интегралу от рациональной дроби его приводит подстановка $z = \left(\frac{a + bt}{t}\right)^{1/s}$.

Итак, в трех случаях, когда одно из чисел p, q или $p + q$ является целым, интеграл (21.14) при помощи указанных

выше подстановок приводится к интегралу от рациональной дроби.

Применительно к интегралу (21.13) этот результат выглядит следующим образом: когда одно из чисел p, $\frac{m+1}{n}$ или $\frac{m+1}{n}+p$ является целым, интеграл (21.13) может быть сведен к интегралу от рациональной дроби. При этом в том случае, когда p целое, это сведение осуществляет подстановка $z = x^{n/s}$, где число s является знаменателем дроби $\frac{m+1}{n}$, т. е. $\frac{m+1}{n} = \frac{r}{s}$; в том случае, когда $\frac{m+1}{n}$ — целое, подстановка

$$z = (a + bx^n)^{1/s},$$

где число s является знаменателем дроби p, т. е. $p = \frac{r}{s}$, а в том случае, когда $\frac{m+1}{n}+p$ — целое, — подстановка

$$z = (ax^{-n} + b)^{1/s},$$

где число s также является знаменателем дроби p.

Этот факт был известен еще И. Ньютону. Л. Эйлер высказал предположение, что ни для каких других показателей m, n и p интеграл от дифференциального бинома нельзя свести к интегралу от рациональных функций. Для рациональных показателей m, n и p, не удовлетворяющих указанным выше условиям, это было доказано П. Л. Чебышёвым, а для иррациональных — Д. Д. Мордухай-Болотовским[1].

Пример. Рассмотрим интеграл

$$I = \int \sqrt{x}\,\sqrt[4]{1 - 1/\sqrt{x^3}}\,dx = \int x^{1/2}(1 - x^{-3/2})^{1/4}\,dx.$$

Здесь $m = 1/2$, $n = -3/2$, $p = 1/4$ и $(m+1)/n = -1$; имеем второй случай.

Сделаем указанную выше подстановку:

$$z = (1 - x^{-3/2})^{1/4}; \qquad (21.15)$$

[1] Д. Д. Мордухай-Болотовский (1876—1952) — русский математик.

отсюда $x = (1 - z^4)^{-2/3}$, $dx = \frac{8}{3}(1 - z^4)^{-5/3} z^3\, dz$, поэтому

$$I = \frac{8}{3}\int \frac{z^4}{(1 - z^4)^2}\, dz = \frac{2}{3}\int z\, d\frac{1}{1 - z^4} = \frac{2}{3}\left(\frac{z}{1 - z^4} - \int \frac{dz}{1 - z^4}\right) =$$

$$= \frac{2z}{3(1 - z^4)} - \frac{1}{3}\int\left(\frac{1}{1 - z^2} + \frac{1}{1 + z^2}\right) dz =$$

$$= \frac{2z}{3(1 - z^4)} - \frac{1}{6}\ln\left|\frac{1 + z}{1 - z}\right| - \frac{1}{3}\operatorname{arctg} z + C,$$

где z выражается через x по формуле (21.15).

21.5. Интегралы вида $\int \frac{P_n(x)dx}{\sqrt{ax^2 + bx + c}}$

Рассмотрим интеграл

$$\int \frac{P_n(x)}{\sqrt{ax^2 + bx + c}}\, dx, \quad a \neq 0,$$

где $P_n(x)$ — многочлен степени $n \geqslant 1$. С принципиальной точки зрения этот интеграл всегда можно свести к интегралу от рациональной дроби с помощью одной из подстановок Эйлера (см. п. 21.3). Однако в данном конкретном случае значительно быстрее к цели приводит обычно другой прием.

Именно, покажем, что справедлива формула

$$\int \frac{P_n(x)}{\sqrt{ax^2 + bx + c}}\, dx =$$

$$= P_{n-1}(x)\sqrt{ax^2 + bx + c} + \alpha\int \frac{dx}{\sqrt{ax^2 + bx + c}}, \qquad (21.16)$$

где $P_{n-1}(x)$ — многочлен степени не выше, чем $n - 1$, а α — некоторое число.

Итак, пусть многочлен

$$P_n(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0 \qquad (21.17)$$

задан. Если существует многочлен

$$P_{n-1}(x) = b_{n-1}x^{n-1} + b_{n-2}x^{n-2} + \ldots + b_0, \qquad (21.18)$$

удовлетворяющий условию (21.16), то, дифференцируя это равенство, получим

$$\frac{P_n(x)}{\sqrt{ax^2+bx+c}} =$$

$$= P'_{n-1}(x)\sqrt{ax^2+bx+c} + \frac{P_{n-1}(x)(2ax+b)}{2\sqrt{ax^2+bx+c}} + \frac{\alpha}{\sqrt{ax^2+bx+c}},$$

или

$$2P_n(x) = 2P'_{n-1}(x)(ax^2+bx+c) + P_{n-1}(x)(2ax+b) + 2\alpha. \quad (21.19)$$

Здесь слева стоит многочлен степени n, а справа каждое слагаемое также является многочленом степени не больше n.

Замечая, что

$$P'_{n-1}(x) =$$

$$= (n-1)b_{n-1}x^{n-2} + \ldots + k\, b_k x^{k-1} + \ldots + b_1, \quad (21.20)$$

и подставляя (21.17), (21.18) и (21.20) в (21.19), имеем равенства

$$2(a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0) =$$

$$= 2(ax^2+bx+c)[(n-1)b_{n-1}x^{n-2} + \ldots + kb_k x^{k-1} + \ldots + b_1] +$$

$$+ (2ax+b)(b_{n-1}x^{n-1} + \ldots + b_k x^k + \ldots + b_0) + 2\alpha.$$

Приравнивая коэффициенты у одинаковых степеней x, получим следующую систему $n+1$ линейных уравнений с $n+1$ неизвестными $b_0, b_1, \ldots, b_{n-1}, \alpha$:

$$2a_0 = 2cb_1 + bb_0 + 2\alpha,$$

$$2a_1 = 2bb_1 + 4cb_2 + 2ab_0 + bb_1,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$2a_k = 2(k-1)ab_{k-1} + 2kbb_k + 2(k+1)cb_{k+1} + 2ab_{k-1} + bb_k,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$2a_{n-1} = 2(n-2)ab_{n-2} + 2(n-1)bb_{n-1} + 2ab_{n-2} + bb_{n-1},$$

$$2a_n = 2(n-1)ab_{n-1} + 2ab_{n-1}. \quad (21.21)$$

Из последнего уравнения сразу находим: $b_{n-1} = a_n/na$. Подставляя это выражение в предпоследнее уравнение и замечая, что в этом уравнении коэффициент у неизвестного b_{n-2} равен $2a(n-1) \neq 0$, найдем значение b_{n-2}. Подставляя далее значения b_{n-1} и b_{n-2} в предыдущее уравнение, найдем

значение b_{n-3}, и т. д. Последовательно получим все значения неизвестных b_k ($k = 0, 1, \ldots, n - 1$). После этого из первого уравнения сразу находится неизвестное α.

Таким образом, система (21.21) имеет решение при любых значениях $a_0, a_1, \ldots, a_n$, поэтому определитель этой системы не равен нулю и указанное решение единственно.

На практике многочлен $P_{n-1}(x)$ в формуле (21.16) пишут с неопределенными коэффициентами, которые находят, решая систему (21.21). Затем вычисление данного интеграла сводится к вычислению интеграла $\int \frac{dx}{\sqrt{ax^2 + bx + c}}$, который в случае, когда подкоренное выражение положительно на некотором промежутке, легко сводится к табличному (см. п. 18.3).

Интегралы вида $\int \frac{dx}{(x - \lambda)^k \sqrt{ax^2 + bx + c}}$ подстановкой $t = \frac{1}{x - \lambda}$ сводятся к интегралам рассмотренного типа (21.16).

§ 22.

Интегрирование некоторых трансцендентных функций

22.1. Интегралы вида $\int R(\sin x, \cos x)\, dx$

Подстановка $u = \operatorname{tg} \frac{x}{2}$, $-\pi < x < \pi$, сводит интегралы вида $\int R(\sin x, \cos x)\, dx$ к интегралу от рациональной дроби. Действительно, имеем

$$\sin x = \frac{2\sin\frac{x}{2}\cos\frac{x}{2}}{\cos^2\frac{x}{2} + \sin^2\frac{x}{2}} = \frac{2\operatorname{tg}\frac{x}{2}}{1 + \operatorname{tg}^2\frac{x}{2}} = \frac{2u}{1 + u^2}, \qquad (22.1)$$

$$\cos x = \frac{\cos^2\frac{x}{2} - \sin^2\frac{x}{2}}{\cos^2\frac{x}{2} + \sin^2\frac{x}{2}} = \frac{1 - \operatorname{tg}^2 x}{1 + \operatorname{tg}^2 x} = \frac{1 - u^2}{1 + u^2},$$

$$x = 2\operatorname{arctg} u, \quad dx = \frac{2du}{1 + u^2},$$

поэтому

$$\int R(\sin x, \cos x)\, dx = 2\int R\left(\frac{2u}{1 + u^2}, \frac{1 - u^2}{1 + u^2}\right)\frac{du}{1 + u^2}.$$

Таким образом получен интеграл от рациональной функции.

Вычислим указанным методом интеграл $\int\frac{dx}{1+\sin x}$. Используя формулы (22.1), получим

$$\int\frac{dx}{1+\sin x}=2\int\frac{du}{(1+u)^2}=-\frac{2}{1+u}+C=-\frac{2}{1+\operatorname{tg}\frac{x}{2}}+C.$$

Следует, однако, иметь в виду, что, хотя с принципиальной точки зрения рассматриваемые интегралы всегда можно привести к интегралу от рациональной дроби указанным методом, при практическом его применении он часто приводит к громоздким вычислениям; вместе с тем другие методы, в частности подстановки вида

$$u=\sin x,\quad u=\cos x,\quad u=\operatorname{tg} x, \tag{22.2}$$

иногда значительно быстрее позволяют вычислить нужный интеграл.

Примеры. 1. Рассмотрим интеграл $\int\frac{dx}{\cos^4 x}$. Представим его в виде $\int\frac{dx}{\cos^4 x}=\int\frac{1}{\cos^2 x}\frac{dx}{\cos^2 x}$. Сразу видно, что в этом случае очень удобна подстановка $u=\operatorname{tg} x$:

$$\int\frac{dx}{\cos^4 x}=\int(1+\operatorname{tg}^2 x)\,d\operatorname{tg} x=\int(1+u^2)\,du=$$

$$=u+\frac{u^3}{3}+C=\operatorname{tg} x+\frac{\operatorname{tg}^3 x}{3}+C.$$

2. Представляя интеграл $\int\frac{dx}{\sin^3 x\cos x}$ в виде $\int\frac{dx}{\sin^3 x\cos x}=$ $=\int\frac{\sin x\,dx}{\sin^4 x\cos x}$, убеждаемся в целесообразности подстановки $u=\cos x$. Действительно,

$$\int\frac{dx}{\sin^3 x\cos x}=\int\frac{\sin x\,dx}{\sin^4 x\cos x}=-\int\frac{d\cos x}{\sin^4 x\cos x}=$$

$$=-\int\frac{du}{(1-u^2)^2u}=-\frac{1}{2}\int\frac{du^2}{(1-u^2)^2u^2}=-\frac{1}{2}\int\frac{dv}{(1-v)^2v}{}^{1}=-\frac{1}{2}\int\frac{(1-v)+v}{(1-v)^2v}\,dv=$$

$$=-\frac{1}{2}\int\frac{dv}{(1-v)v}-\frac{1}{2}\int\frac{dv}{(1-v)^2}=-\frac{1}{2}\int\frac{(1-v)+v}{(1-v)v}\,dv-\frac{1}{2}\,\frac{1}{1-v}=$$

$$=\frac{1}{2}\int\frac{dv}{v}-\frac{1}{2}\int\frac{dv}{1-v}-\frac{1}{2}\,\frac{1}{1-v}=-\frac{1}{2}\ln|v|+\frac{1}{2}\ln|1-v|-\frac{1}{2(1-v)}+C=$$

$$=\ln|\operatorname{tg} x|-\frac{1}{2\sin^2 x}+C.$$

---

[1] Здесь сделана подстановка $v=u^2$.

Конечно, интегралы, рассмотренные в примерах 1 и 2, могут быть вычислены и с помощью подстановки (22.1), например,

$$\int \frac{dx}{\sin^3 x \cos x} = \frac{1}{4}\int \frac{(1+u^2)^3 du}{u^3(1-u^2)};$$

однако при таком способе пришлось бы интегрировать более сложную рациональную дробь, чем в результате применения подстановки $u = \cos x$.

3. Иногда при вычислении интегралов, подынтегральное выражение которых содержит $\sin x$ и $\cos x$, бывает полезно прибегать и к другим искусственным приемам, используя известные тригонометрические формулы, как, например, формулу $\sin^2 x + \cos^2 x = 1$. Покажем на рассмотренном только что примере способ применения этой формулы:

$$\int \frac{dx}{\sin^3 x \cos x} = \int \frac{\sin^2 x + \cos^2 x}{\sin^3 x \cos x} dx = \int \frac{dx}{\sin x \cos x} + \int \frac{\cos x dx}{\sin^3 x} =$$

$$= \int \frac{d\operatorname{tg} x}{\operatorname{tg} x} + \int \frac{d\sin x}{\sin^3 x} = \int \frac{du}{u} + \int \frac{dv}{v^3} = \ln|u| - \frac{1}{2v^2} + C =$$

$$= \ln|\operatorname{tg} x| - \frac{1}{2\sin^2 x} + C.$$

Как и следовало ожидать, получился тот же результат, что и выше.

22.2. Интегралы вида $\int \sin^m x \cos^n dx$

Пусть m и n — рациональные числа. Интеграл $\int \sin^m x \cos^n dx$ с помощью подстановок $u = \sin x$ или $u = \cos x$ сводится к интегралу от дифференциального бинома.

Действительно, полагая, например, $u = \sin x$, получаем

$$\cos x = \sqrt{1-\sin^2 x} = (1-u^2)^{1/2},\ du = \cos x\, dx,$$

$$dx = \frac{du}{\cos x} = (1-u^2)^{-1/2} du,$$

поэтому

$$\int \sin^m x \cos^n x\, dx = \int u^m (1-u^2)^{(n-1)/2}\, du.$$

Таким образом, можно выразить или нет интеграл $\int \sin^m x \cos^n x \; dx$ через элементарные функции, зависит от того, обладает этим свойством или нет получающийся интеграл от дифференциального бинома.

В том случае, когда m и n — целые (не обязательно положительные) числа, интеграл $\int \sin^m x \cos^n x \, dx$ относится к типу интегралов, рассмотренных в предыдущем пункте, в частности, для их вычисления целесообразно применять подстановки (22.2).

Например, когда $m = 2k + 1$ (соответственно $n = 2k + 1$) — нечетное число, то можно сделать подстановку $u = \cos x$ (соответственно $u = \sin x$):

$$\int \sin^{2k+1} x \cos^n x \, dx = -\int (1 - \cos^2 x)^k \cos^n x \, d\cos x =$$
$$= -\int (1 - u^2)^k u^n \, du.$$

Рассматриваемый интеграл сведен к интегралу от рациональной дроби.

Аналогичный результат можно получить и для интеграла $\int \sin^m x \cos^{2k+1} x \, dx$ с помощью подстановки $u = \sin x$.

Если $m = 2k + 1$, $n = 2l + 1$, то бывает полезной подстановка $t = \cos 2x$:

$$\int \sin^{2k+1} x \cos^{2l+1} dx = \int \sin^{2k} x \cos^{2l} x \sin x \cos x \, dx =$$
$$= \frac{1}{2} \int \sin^{2k} x \cos^{2l} x \sin 2x \, dx =$$
$$= -\frac{1}{4} \int \left(\frac{1 - \cos 2x}{2}\right)^k \left(\frac{1 + \cos 2x}{2}\right)^l d \cos 2x =$$
$$= \frac{1}{2^{k+l+2}} \int (1 - t)^k (1 + t)^l \, dt,$$

т. е. снова получен интеграл от рациональной дроби; подчеркнем, что здесь k и l могут быть как положительными, так и отрицательными целыми числами.

Если оба показателя m и n положительны и четны (или один из них нуль), то целесообразно применять формулы

$$\sin^2 x = \frac{1 - \cos 2x}{2}, \; \cos^2 x = \frac{1 + \cos 2x}{2},$$

которые, очевидно, приводят рассматриваемый интеграл к интегралам того же типа, но с меньшими, также неотрицательными показателями. Например,

$$\int \cos^2 x \, dx = \int \frac{1 + \cos 2x}{2} \, dx = \frac{x}{2} + \frac{\sin 2x}{4} + C.$$

22.3. Интегралы вида $\int \sin \alpha x \cos \beta x \, dx$

Интегралы

$$\int \sin \alpha x \cos \beta x \, dx, \quad \int \sin \alpha x \sin \beta x \, dx \quad \text{и} \quad \int \cos \alpha x \cos \beta x \, dx$$

непосредственно вычисляются, если их подынтегральные функции преобразовать по формулам

$$\sin \alpha x \cos \beta x = \frac{1}{2}[\sin(\alpha + \beta)x + \sin(\alpha - \beta)x],$$

$$\sin \alpha x \sin \beta x = \frac{1}{2}[\cos(\alpha - \beta)x - \cos(\alpha + \beta)x],$$

$$\cos \alpha x \cos \beta x = \frac{1}{2}[\cos(\alpha + \beta)x + \cos(\alpha - \beta)x].$$

Например,

$$\int \sin 2x \cos x \, dx = \frac{1}{2}\int (\sin 3x + \sin x)\, dx =$$

$$= -\frac{1}{6}\cos 3x - \frac{1}{2}\cos x + C.$$

22.4. Интегралы от трансцендентных функций, вычисляющиеся с помощью интегрирования по частям

К таким интегралам относятся, например, интегралы

$$\int e^{\alpha x}\cos \beta x \, dx, \int e^{\alpha x}\sin \beta x \, dx, \int x^n \cos \alpha x \, dx, \int x^n \sin \alpha x \, dx,$$
$$\int x^n e^{\alpha x} \, dx, \int x^n \arcsin x \, dx, \int x^n \arccos x \, dx, \int x^n \operatorname{arctg} x \, dx,$$
$$\int x^n \operatorname{arcctg} x \, dx, \int x^n \ln x \, dx \ (n \text{ — целое неотрицательное}).$$

Все эти интегралы вычисляются с помощью, вообще говоря, последовательного интегрирования по частям. Действительно, имеем

$$I = \int e^{\alpha x}\cos \beta x \, dx = \int e^{\alpha x}\, d\frac{\sin \beta x}{\beta} =$$

$$= \frac{e^{\alpha x}\sin \beta x}{\beta} - \frac{\alpha}{\beta}\int e^{\alpha x}\sin \beta x \, dx = \frac{e^{\alpha x}\sin \beta x}{\beta} - \frac{\alpha}{\beta}\int e^{\alpha x}\, d\left(-\frac{\cos \beta x}{\beta}\right) =$$

$$= \frac{e^{\alpha x}\sin \beta x}{\beta} + \frac{\alpha e^{\alpha x}\cos \beta x}{\beta^2} - \frac{\alpha^2}{\beta^2}\int e^{\alpha x}\cos \beta x \, dx =$$

$$= \frac{e^{\alpha x}(\beta \sin \beta x + \alpha \cos \beta x)}{\beta^2} - \frac{\alpha^2}{\beta^2} I,$$

откуда

$$I = \frac{e^{\alpha x}(\beta \sin \beta x + \alpha \cos \beta x)}{\alpha^2 + \beta^2} + C. \qquad (22.3)$$

Аналогично интегралу $\int e^{\alpha x} \cos \beta x \, dx$ вычисляется интеграл

$$\int e^{\alpha x} \sin \beta x \, dx = \frac{e^{\alpha x}(\alpha \sin \beta x + \beta \cos \beta x)}{\alpha^2 + \beta^2},$$

а через эти два интеграла легко выражаются интегралы

$$\int \operatorname{sh} \alpha x \cos \beta x \, dx, \quad \int \operatorname{sh} \alpha \sin \beta x \, dx,$$
$$\int \operatorname{ch} \alpha x \cos \beta x \, dx, \quad \int \operatorname{ch} \alpha x \sin \beta x \, dx.$$

Впрочем, эти последние четыре интеграла можно вычислить и непосредственно с помощью интегрирования по частям подобно тому, как был вычислен рассмотренный выше интеграл (22.3).

В интегралах $\int x^n \cos \alpha \, dx$, $\int x^n \sin \alpha x \, dx$, $\int x^n e^{\alpha x} \, dx$, положив $u = x^n$ и соответственно $dv = \cos \alpha x \, dx$, $dv = \sin \alpha x \, dx$, $dv = e^{\alpha x} \, dx$, после интегрирования по частям снова придем к интегралу одного из указанных видов, но уже с показателем степени у переменной x, меньшим на единицу. Применяя этот прием n раз, придем к интегралу рассматриваемого типа с $n = 0$, который, очевидно, сразу берется. Например,

$$\int x^2 \sin x \, dx = \int x^2 \, d(-\cos x) = -x^2 \cos x + 2\int x \cos x \, dx =$$
$$= -x^2 \cos x + 2\int x \, d\sin x = -x^2 \cos x + 2x \sin x - 2\int \sin x \, dx =$$
$$= -x^2 \cos x + 2x \sin x + 2\cos x + C.$$

Используя интегралы, рассмотренные выше, можно вычислить и более сложные интегралы. Вычислим, например, интеграл $\int x^n e^{\alpha x} \cos \beta x \, dx$.

Интегрируя по частям и применяя (22.3), имеем

$$\int x^n e^{\alpha x} \cos \beta x \, dx = \int x^n d\left[\frac{e^{\alpha x}(\beta \sin \beta x + \alpha \cos \beta x)}{\alpha^2 + \beta^2}\right] =$$
$$= x^n e^{\alpha x} \frac{\beta \sin \beta x + \alpha \cos \beta x}{\alpha^2 + \beta^2} - \frac{n\beta}{\alpha^2 + \beta^2} \int x^{n-1} e^{\alpha x} \sin \beta x \, dx -$$
$$- \frac{n\alpha}{\alpha^2 + \beta^2} \int x^{n-1} e^{\alpha x} \cos \beta x \, dx.$$

В правой части равенства получились интегралы того же типа, что и исходный, только степень у x на единицу меньше. Применяя последовательно указанный прием, придем к интегралам вида $\int e^{\alpha x}\sin \beta x\, dx$ и $\int e^{\alpha x}\cos \beta x\, dx$, которые рассмотрены выше.

Наконец, интегралы

$$\int x^n \arcsin x\, dx,\ \int x^n \arccos x\, dx,\ \int x^n \operatorname{arcctg} x\, dx \text{ и } \int x^n \ln x\, dx$$

сводятся интегрированием по частям к интегралу от алгебраической функции, если в них положить $dv = x^n\, dx$, а за функцию u взять оставшуюся трансцендентную функцию, т. е. одну из функций $\arcsin x$, $\arccos x$, $\operatorname{arctg} x$, $\operatorname{arcctg} x$, $\ln x$. Например,

$$\int x \ln x\, dx = \int \ln x\, d\frac{x^2}{2} =$$
$$= \frac{x^2 \ln x}{2} - \frac{1}{2}\int x\, dx = \frac{x^2 \ln x}{2} - \frac{x^2}{4} + C.$$

22.5. Интегралы вида $\int R(\operatorname{sh} x,\ \operatorname{ch} x)\, dx$

Подстановка $u = \operatorname{th}\frac{x}{2}$ сводит интеграл $\int R(\operatorname{sh} x,\ \operatorname{ch} x)\, dx$ к интегралу от рациональной дроби.

Действительно, при указанной замене переменной имеем

$$\operatorname{sh} x = \frac{2u}{1-u^2}, \quad \operatorname{ch} x = \frac{1+u^2}{1-u^2}, \quad dx = \frac{2du}{1-u^2},$$

поэтому

$$\int R(\operatorname{sh} x,\ \operatorname{ch} x)\, dx = 2\int R\left(\frac{2u}{1-u^2},\ \frac{1+u^2}{1-u^2}\right)\frac{du}{1-u^2}.$$

В конкретных примерах иногда оказывается значительно удобнее использовать подстановки вида $u = \operatorname{sh} x$, $u = \operatorname{ch} x$ или $u = \operatorname{tg} x$, позволяющие вычислить интеграл существенно проще (ср. п. 22.1).

Интегралы вида $\int \operatorname{sh}^m x \operatorname{ch}^n x\, dx$, где m и n — рациональные числа, с помощью подстановок $v = \operatorname{sh} x$ ($u = \operatorname{ch} x$) приводятся к интегралу от дифференциального бинома (ср. п. 22.2).

22.6. Замечания об интегралах, не выражающихся через элементарные функции

Мы рассмотрели различные классы элементарных функций и нашли их первообразные, которые также являются элементарными функциями. Однако не всякая элементарная

функция имеет в качестве своей первообразной элементарную же функцию. С подобным обстоятельством мы уже встретились при рассмотрении интеграла от дифференциального бинома: в этом случае подынтегральная функция — элементарная (иррациональная), а интеграл от нее, как отмечалось, вычисляется далеко не всегда.

Можно показать, что интегралы $\int \frac{e^x}{x^n}\,dx$, $\int \frac{\sin x}{x^n}\,dx$, $\int \frac{\cos x}{x^n}\,dx$ (n — натуральное число) также не выражаются через элементарные функции.

Имеется ряд интегралов от элементарных функций, не выражающихся через элементарные функции и играющих большую роль как в самом математическом анализе, так и в его разнообразных приложениях. К таким интегралам относится, например, интеграл $\int e^{-x^2}\,dx$, а также так называемые *эллиптические интегралы* $\int R\,(x, \sqrt{P(x)})\,dx$, где $P(x)$ — многочлен третьей или четвертой степени. В общем случае эти интегралы не выражаются через элементарные функции. Особенно часто встречаются интегралы

$$\int \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} \quad \text{и} \quad \int \frac{x^2dx}{\sqrt{(1-x^2)(1-k^2x^2)}}, \quad 0 < k < 1,$$

которые подстановкой $x = \sin\varphi$ приводятся к линейным комбинациям интегралов

$$\int \frac{d\varphi}{\sqrt{1-k^2\sin^2\varphi}} \quad \text{и} \quad \int \sqrt{1-k^2\sin^2\varphi}\;d\varphi,$$

они называются соответственно *эллиптическими интегралами первого и второго рода в форме Лежандра.*

§ 23.

Определенный интеграл

23.1. Определение интеграла Римана

Напомним (см. п. 16.5), что *разбиением* τ отрезка $[a, b]$ называется любая конечная система его точек x_i, $i = 0, 1, 2, \ldots, i_\tau$, такая, что

$$a = x_0 < x_1 < \ldots < x_{i_\tau - 1} < x_{i_t} = b.$$

При этом пишут $\tau = \{x_i\}_{i=0}^{i=i_\tau}$. Каждый из отрезков $[x_{i-1}, x_i]$ называется *отрезком разбиения* τ, его длину обозначают через Δx_i, $\Delta x_i = x_i - x_{i-1}$, $i = 1, 2, \ldots, i_\tau$.

Число $|\tau| = \max\limits_{i=1,2,\ldots,i_\tau} \Delta x_i$ называется *мелкостью разбиения* τ.

Разбиение τ' отрезка $[a, b]$ называется следующим за разбиением τ (или продолжающим разбиение τ) того же отрезка, а также вписанным в разбиение τ, если каждая точка разбиения τ является и точкой разбиения τ', иначе говоря, если каждый отрезок разбиения τ' содержится в некотором отрезке разбиения τ (говорят еще, что τ' — измельчение разбиения τ). В этом случае пишут $\tau' \succ \tau$ или, что то же, $\tau \prec \tau'$.

Совокупность всех разбиений данного отрезка обладает следующими свойствами.

1^0. *Если* $\tau_1 \prec \tau_2$, *а* $\tau_2 \prec \tau_3$, *то* $\tau_1 \prec \tau_3$.

2^0. *Для любых* τ_1 *и* τ_2 *существует такое* τ, *что* $\tau \succ \tau_1$ *и* $\tau \succ \tau_2$.

В самом деле, первое свойство следует просто из того, что, в силу условия $\tau_3 \succ \tau_2$, каждый отрезок разбиения τ_3 содержится в некотором отрезке разбиения τ_2, который, в свою очередь, согласно условию $\tau_2 \succ \tau_1$, содержится в каком-то отрезке разбиения τ_1; таким образом, всякий отрезок разбиения τ_3 лежит на определенном отрезке разбиения τ_1, а это и означает, что $\tau_3 \succ \tau_1$.

Для доказательства свойства 2^0 разбиений заметим лишь, что если заданы два разбиения τ_1 и τ_2, то разбиение τ, состоящее из всех точек, входящих как в разбиение τ_1, так и в разбиение τ_2, очевидно, будет следовать за τ_1 и за τ_2. □

Пусть теперь на отрезке $[a, b]$ определена функция f и пусть $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ — некоторое разбиение этого отрезка, $\Delta x_i = x_i - x_{i-1}$, $i = 1, 2, \ldots, i_\tau$, а $|\tau|$ — мелкость этого разбиения.

Зафиксируем произвольным образом точки $\xi_i \in [x_{i-1}, x_i]$, $i = 1, 2, \ldots, i_\tau$, и составим сумму:

$$\sigma_\tau(f; \xi_1, \ldots, \xi_{i_\tau}) = \sum_{i=1}^{i_\tau} f(\xi_i)\Delta x_i .$$

Суммы вида $\sigma_\tau(f; \xi_1, \ldots, \xi_{i_\tau})$ называются *интегральными суммами Римана функции* f. Иногда для краткости будем их обозначать через $\sigma_\tau(f)$, или даже просто через σ_τ. Геометрически в том случае, когда функция f неотрицательна (рис. 105), каждое слагаемое интегральной суммы Римана σ_τ равно площади прямоугольника с основанием длины Δx_i и с высотой $f(\xi_i)$. Вся же сумма σ_τ равна площади «ступенчатой фигуры», получающейся объединением всех указанных прямоугольников.

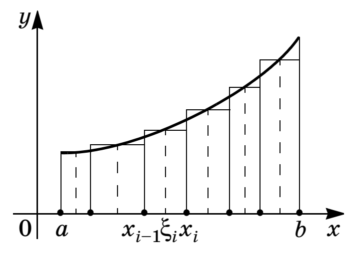


Рис. 105

Определение 1. *Функция f называется интегрируемой (по Риману) на отрезке $[a, b]$, если существует такое число A, что для любой последовательности разбиений отрезка $[a, b]$*

$$\tau_n = \{x_i^{(n)}\}_{i=0}^{i=i_{\tau_n}}, \quad n = 1, 2, \ldots,$$

у которой $\lim\limits_{n\to\infty}|\tau_n| = 0$, *и для любого выбора точек*

$$\xi_i^{(n)} \in [x_{i-1}^{(n)}, x_i^{(n)}], \quad i = 1, 2, \ldots, i_{\tau_n}, \; n = 1, 2, \ldots,$$

существует предел последовательности интегральных сумм $\sigma_{\tau_n}(f; \xi_1^{(n)}, \ldots, \xi_{i_{\tau_n}}^{(n)})$ *и он равен* A:

$$\lim_{n\to\infty} \sum_{i=1}^{i_{\tau_n}} f(\xi_i^{(n)})\Delta x_i^{(n)} = A, \tag{23.1}$$

где $\Delta x_i^{(n)} = x_i^{(n)} - x_{i-1}^{(n)}, \quad i = 1, 2, \ldots, i_{\tau_n}, \; n = 1, 2, \ldots$.

При выполнении этих условий число A называется (римановым) определенным интегралом функции f на отрезке $[a, b]$ и обозначается $\int\limits_a^b f(x)\,dx$.

Выражение $\int\limits_a^b f(x)\,dx$ читается: «интеграл от a до b $f(x)\,dx$»; x называется *переменной интегрирования*, f — *подынтегральной функцией*, a — *нижним*, а b — *верхним пределом интеграла*, или, соответственно, верхним и нижним пределом интегрирования; *отрезок $[a, b]$ — промежутком интегрирования*.

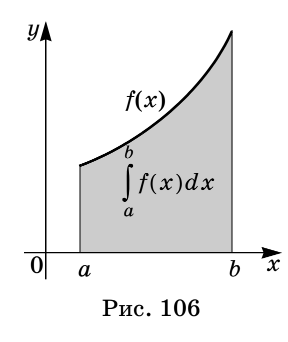


Рис. 106

Таким образом,

$$\int_a^b f(x)\,dx \overset{\text{def}}{=} \lim_{n\to\infty} \sigma_{\tau_n}(f;\, \xi_i^{(n)},\, \ldots,\, \xi_{i_{\tau_n}}^{(n)}),$$

где последовательность τ_n такова, что

$$\lim_{n\to\infty} |\tau_n| = 0.$$

Для краткости будем в этом случае писать просто

$$\int_a^b f(x)\,dx = \lim_{|\tau|\to 0} \sigma_\tau(f).$$

Подобно тому как определение предела функции можно сформулировать двумя эквивалентными способами с помощью пределов последовательностей и с помощью «(ε—δ)-языка», так и определение интеграла можно сформулировать иначе.

Определение 2. *Число A называется определенным интегралом функции f на отрезке* $[a, b]$, *если для любого* $\varepsilon > 0$ *существует такое* $\delta > 0$, *что каково бы ни было разбиение* $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ *(отрезка* $[a, b]$*) мелкости, меньшей* δ: $|\tau| < \delta$, *и каковы бы ни были точки* $\xi_i \in [x_{i-1}, x_i]$, *выполняется неравенство*

$$\left| \sum_{i=1}^{i_\tau} f(\xi_i)(\Delta x_i) - A \right| < \varepsilon,$$

где $\Delta x_i = x_i - x_{i-1}$, $i = 1, 2, \ldots, i_\tau$.

УПРАЖНЕНИЕ 1. Докажите, что два данных выше определения определенного интеграла эквивалентны.

Из определения 1 следует, что для неотрицательных функций определенный интеграл $\int_a^b f(x)\,dx$ является пределом при $|\tau| \to 0$ последовательности площадей соответствующих ступенчатых фигур; поэтому он, естественно, оказывается связанным с понятием площади, а именно равен площади фигуры[1] (называемой «криволинейной трапецией»), границей которой является график функции f, отрезок $[a, b]$ оси Ox и, быть может, отрезки прямых $x = a$ и $x = b$,

[1] Привычный из элементарной геометрии термин «фигура» употребляется здесь всюду в смысле «плоское множество».

ординаты точек которых меняются соответственно от нуля до $f(a)$ и до $f(b)$ (рис. 106). Для того чтобы это доказать, надо прежде всего уточнить само понятие площади рассматриваемых фигур. Все это будет сделано ниже, в § 27.

Замечание. Интегральные суммы являются функциями, аргументами которых являются разбиения τ и точки ξ_i, $i = 1, 2, \ldots, i_\tau$, лежащие на отрезках этих разбиений. Однако предел интегральных сумм — это не предел функции в ранее определенном смысле, а принципиально новое понятие. При переходе к пределу интегральных сумм их аргументы τ и ξ_i, $i = 1, 2, \ldots, i_\tau$, изменяются, но ни к чему не стремятся. Здесь нельзя, как в случае предела функции, рассмотренного ранее, указать «точку» $(\tau, \xi_1, \xi_2, \ldots, \xi_{i_\tau})$, в которой берется предел интегральных сумм. Следовательно, введенное здесь понятие предела интегральных сумм Римана является новым понятием, не укладывающимся ни в понятие предела последовательности, ни в понятие предела функции.

В дальнейшем придется использовать аналогичное понятие предела не только для интегральных сумм Римана, но и для других объектов. Поэтому сформулируем общее определение предела этого вида.

Определение 3. *Рассмотрим множество* $\Gamma = \{\tau\}$ *всех разбиений отрезка* $[a, b]$. *Пусть на этом множестве определена числовая, вообще говоря, многозначная функция* $\Phi(\tau)$, $\tau \in \Gamma$. *Будем говорить, что функция* $\Phi(\tau)$ *при* $|\tau| \to 0$ *имеет предел, равный* A, *и писать*

$$\lim_{|\tau| \to 0} \Phi(\tau) = A,$$

если для любой последовательности разбиений $\tau_n \in \Gamma$, $n = 1, 2, \ldots$, *такой, что* $\lim\limits_{n \to \infty} |\tau_n| = 0$, *при любом выборе значений* $\Phi(\tau_n)$ *числовая последовательность* $\Phi(\tau_n)$ *сходится к числу* A, *т. е.*

$$\lim_{n \to \infty} \Phi(\tau_n) = A.$$

Это понятие предела определено с помощью понятия предела последовательности, поэтому для него справедливы

многие свойства, аналогичные соответствующим свойствам предела последовательности.

Пусть, например, $\Phi_1(\tau)$ и $\Phi_2(\tau)$ — однозначные функции разбиений τ некоторого отрезка $[a, b]$ и существуют пределы $\lim\limits_{|\tau| \to 0} \Phi_1(\tau)$ и $\lim\limits_{|\tau| \to 0} \Phi_2(\tau)$. Тогда если для всех τ выполняется неравенство $\Phi_1(\tau) \leqslant \Phi_2(\tau)$, то и $\lim\limits_{|\tau| \to 0} \Phi_1(\tau) \leqslant \lim\limits_{|\tau| \to 0} \Phi_2(\tau)$. А если $\Psi(\tau)$ — такая функция, что для всех τ имеют место неравенства

$$\Phi_1(\tau) \leqslant \Psi(\tau) \leqslant \Phi_2(\tau)$$

и

$$\lim_{|\tau| \to 0} \Phi_1(\tau) = \lim_{|\tau| \to 0} \Phi_2(\tau) = a,$$

то существует предел $\lim\limits_{|\tau| \to 0} \Psi(\tau)$ и он также равен a. Мы будем использовать эти свойства в дальнейшем.

Как и в случае предела интегральных сумм Римана, понятие этого предела можно сформулировать на «ε—δ-языке», что предоставляется читателю.

Отметим в заключение, что многозначность функции Φ, о которой идет речь в определении 3, в случае интегральных сумм Римана связана с различным способом выбора точек $\xi_i \in [x_{i-1}, x_i]$, $i = 1, 2, \ldots, i_\tau$.

Нами было введено понятие определенного интеграла $\int\limits_a^b f(x)\,dx$ от функции f по отрезку $[a, b]$, $a < b$.

Для любой функции f, определенной в точке a, положим, по определению,

$$\int_a^a f(x)\,dx = 0, \tag{23.2}$$

а для функции f, интегрируемой на отрезке $[a, b]$,

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx,\ a < b. \tag{23.3}$$

Эти определения в известной мере естественны. В первом случае, т. е. при $a = b$, естественно, все промежутки разбиения отрезка $[a, b]$ становятся точками, а их длины Δx_i равны нулю. Поэтому все интегральные суммы $\sum\limits_{i=1}^{i_\tau} f(\xi_i)\Delta x_i$ в

этом случае также равны нулю, а вместе с ними обращается в нуль и интеграл в левой части равенства (23.2).

Во втором случае следует считать отрезки $[x_{i-1}, x_i]$ разбиения $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ отрезка $[a, b]$ ориентированными в отрицательном направлении оси Ox (понятие ориентированного отрезка знакомо читателю из аналитической геометрии), поэтому их длины Δx_i — отрицательными. Отсюда следует, что все интегральные суммы, образуемые для интеграла $\int\limits_b^a f(x)\,dx$, отличаются лишь знаком от соответствующих интегральных сумм интеграла $\int\limits_a^b f(x)\,dx$, что и делает естественной формулу (23.3).

Этим интуитивным соображениям можно, конечно, придать и строгую логическую форму, введя соответствующие определения, однако гораздо проще и короче ввести равенства (23.2) и (23.3) по определению.

23.2*. Критерий Коши существования интеграла

Аналогично критерию Коши существования предела функций формулируется и доказывается аналогичный критерий существования предела интегральных сумм.

ТЕОРЕМА 1. *Для того чтобы функция f была интегрируема на отрезке $[a, b]$, необходимо и достаточно, чтобы для любого $\varepsilon > 0$ существовало такое $\delta > 0$, что, каковы бы ни были разбиения $\tau' = \{x_j'\}_{i=0}^{i=i_{\tau'}}$ и $\tau'' = \{x_i''\}_{j=0}^{j=j_{\tau''}}$ мелкостей меньше δ: $|\tau'| < \delta$, $|\tau''| < \delta$ и точки*

$$\xi_i'' \in [x_{i-1}', x_i'],\ i = 1, 2, \ldots, i_{\tau'},$$

$$\xi_j'' \in [x_{j-1}'', x_j''],\ j = 1, 2, \ldots, j_{\tau''},$$

выполняется неравенство

$$|\sigma_{\tau'}(f; \xi_1', \ldots, \xi_{i_{\tau'}}') - \sigma_{\tau''}(f; \xi_1'', \ldots, \xi_{j_{\tau''}}'')| < \varepsilon. \qquad (23.4)$$

Доказательство необходимости условия (23.4). Если функция f интегрируема, т. е. существует предел (23.1), то, согласно определению 2, для любого $\varepsilon > 0$ существу-

ет такое $\delta > 0$, что для любого разбиения $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ мелкости меньше δ: $|\tau| < \delta$ и при любом выборе точек $\xi_i \in [x_{i-1}, x_i]$, $i = 1, 2, \ldots, i_\tau$ для интегральных сумм $\sigma_\tau = \sigma_\tau(f; \xi_1, \ldots, \xi_{i_\tau})$ выполняется неравенство

$$|\sigma_\tau - A| < \frac{\varepsilon}{2}. \qquad (23.5)$$

Если теперь $\sigma_{\tau'} = \sigma_{\tau'}(f; \xi_1', \ldots, \xi_{i'}')$ и $\sigma_{\tau''} = \sigma_{\tau''}(f; \xi_1'', \ldots$ $\ldots, \xi_{j_{\tau''}}'')$ — две такие интегральные суммы, что $|\tau'| < \delta$ и $|\tau''| < \delta$, то

$$|\sigma_{\tau'} - \sigma_{\tau''}| = |(\sigma_{\tau'} - A) + (A - \sigma_{\tau''})| \leqslant$$

$$\leqslant |\sigma_{\tau'} - A| + |A - \sigma_{\tau''}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Доказательство достаточности условия (23.4). Пусть для функции f, заданной на отрезке $[a, b]$, выполнено условие (23.4) и $\sigma_{\tau_n} = \sigma_{\tau_n}(f; \xi_1^{(n)}, \ldots, \xi_{i_{\tau_n}}^{(n)})$, $n = 1, 2, \ldots,$ — такая последовательность интегральных сумм функции f, что

$$\lim_{n \to \infty} |\tau_n| = 0. \qquad (23.6)$$

Если $\varepsilon > 0$ произвольно фиксировано, а $\delta > 0$ выбрано так, что выполняется условие (23.4), то, в силу условия (23.6), существует такой номер n_0, что для всех $n > n_0$ выполняется условие $|\tau_n| < \delta$. Поэтому для любых $n > n_0$ и $m > n_0$ выполняются неравенства $|\tau_n| < \delta$, $|\tau_m| < \delta$ и, следовательно, согласно условию (23.4), имеет место неравенство $|\sigma_{\tau_n} - \sigma_{\tau_m}| < \varepsilon$. Это означает, что числовая последовательность $\{\sigma_{\tau_n}\}$ удовлетворяет критерию Коши сходимости последовательностей и поэтому существует конечный предел $\lim\limits_{n \to \infty} \sigma_{\tau_n}$.

Последовательность $\{\sigma_{\tau_n}\}$ являлась произвольной последовательностью интегральных сумм, для которой выполнялось условие (23.6), поэтому все такие последовательности

сходятся, причем к одному и тому же числу. В самом деле, пусть последовательности $\{\sigma_{\tau'_n}\}$ и $\{\sigma_{\tau''_n}\}$ таковы, что

$$\lim_{n\to\infty}|\tau'_n| = \lim_{n\to\infty}|\tau''_n| = 0 \tag{23.7}$$

и, следовательно, существуют конечные пределы

$$\lim_{n\to\infty}\sigma_{\tau'_n} = A', \quad \lim_{n\to\infty}\sigma_{\tau''_n} = A''. \tag{23.8}$$

Составим новую последовательность интегральных сумм $\{\sigma_{\tau_m}\}$, у которой на нечетных местах стоят члены последовательности $\{\sigma_{\tau'_n}\}$, а на четных — $\{\sigma_{\tau''_n}\}$:

$$\sigma_{\tau_m} = \begin{cases} \sigma_{\tau'_n}, \text{ где } m = 2n-1, \\ \sigma_{\tau''_n}, \text{ где } m = 2n, \end{cases} \qquad n = 1, 2, \dots .$$

Тогда, очевидно, $\left|\lim\limits_{m\to\infty}\right|\tau_m = 0$ и поэтому существует конечный предел

$$\lim_{n\to\infty}\sigma_{\tau_m}.$$

Предел всякой подпоследовательности сходящейся последовательности равен пределу всей последовательности, следовательно,

$$A' = \lim_{n\to\infty}\sigma_{\tau'_n} = \lim_{m\to\infty}\sigma_{\tau_m} = \lim_{n\to\infty}\sigma_{\tau''_n} = A''. \ \square$$

23.3. Ограниченность интегрируемой функции

Установим прежде всего необходимое условие, которому удовлетворяют интегрируемые функции — их ограниченность.

ТЕОРЕМА 2. *Если функция интегрируема на некотором отрезке, то она ограничена на этом отрезке.*

Доказательство. Пусть функция f не ограничена на отрезке $[a, b]$ и пусть фиксировано некоторое разбиение $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ этого отрезка. В силу неограниченности функции f на всем отрезке $[a, b]$, она не ограничена по крайней мере на одном отрезке разбиения τ. Пусть для определенности функция f не ограничена на отрезке $[x_0, x_1]$. Тогда на этом

отрезке существует последовательность $\xi_1^{(n)} \in [x_0, x_1]$, $n = 1, 2, \dots$, такая, что[1]

$$\lim_{n\to\infty} f(\xi_1^{(n)}) = \infty. \tag{23.9}$$

Зафиксируем теперь каким-либо образом точки $\xi_i \in [x_{i-1}, x_i]$, $i = 2, 3, \dots, i_\tau$. Тогда сумма $\sum_{i=2}^{i_\tau} f(\xi_i)\Delta x_i$ будет иметь определенное значение. Поэтому, в силу (23.9),

$$\lim_{n\to\infty} \sigma_\tau(f; \xi_1^{(n)}, \xi_2, \dots, \xi_{i_\tau}) = \lim_{n\to\infty}\left[f'(\xi_1^{(n)})\Delta x_1 + \sum_{i=2}^{i_\tau}(\xi_i)\Delta x_i\right] = \infty$$

и, следовательно, каково бы ни было число $M > 0$, всегда можно подобрать такой номер n_0, что если на первом отрезке $[x_0, x_1]$ взять точку $\xi_i^{(n_0)}$, то

$$|\sigma_\tau(f; \xi_i^{n_0}, \xi_2, \dots, \xi_{i_\tau})| > M.$$

Отсюда следует, что суммы σ_τ не могут стремиться ни к какому конечному пределу при $|\tau| \to 0$.

Действительно, если бы существовал конечный предел $\lim_{|\tau|\to 0} \sigma_\tau = A$, то для любого $\varepsilon > 0$ нашлось бы такое $\delta_\varepsilon > 0$, что для всех разбиений $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ отрезка $[a, b]$ мелкости $|\tau| < \delta_\varepsilon$ при любом выборе точек $\xi_i \in [x_{i-1}, x_i]$, $i = 1, 2, \dots, i_\tau$, выполнялось бы неравенство $|\sigma_\tau - A| < \varepsilon$ и, следовательно,

$$|\sigma_\tau| = |(\sigma_\tau - A) + A| \leqslant |\sigma_\tau - A| + |A| < \varepsilon + |A|.$$

В данном случае, т. е. в случае неограниченности функции f, для любого разбиения τ (в том числе и такого, что $|\tau| < \delta_\varepsilon$, если существовало бы указанное δ_ε) при любом фиксированном $\varepsilon > 0$ можно так выбрать точки ξ_i, что будет

[1] Действительно, в силу неограниченности функции f на отрезке $[x_0, x_1]$, например для любого натурального $n = 1, 2, \dots$, существует такая точка $\xi_1^{(n)} \in [x_0, x_1]$, что $|f(\xi_1^{(n)})| > n$. Очевидно, что последовательность $\{\xi_1^{(n)}\}$ и удовлетворяет условию (23.9).

выполняться неравенство $|\sigma_\tau| > |A| + \varepsilon$. Полученное противоречие доказывает теорему. □

Условие ограниченности функции f, будучи необходимым для ее интегрируемости, не является вместе с тем достаточным. Примером, доказывающим это утверждение, может служить функция Дирихле (см. п. 5.2):

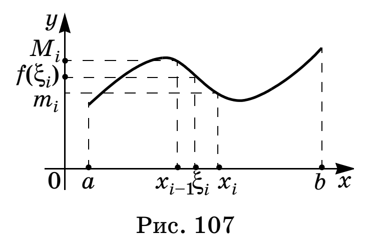


Рис. 107

$$f(x) = \begin{cases} 1, \text{ если } x \text{ рациональное,} \\ 0, \text{ если } x \text{ иррациональное.} \end{cases}$$

Рассмотрим эту функцию на отрезке [0, 1]. Она, очевидно, ограничена на нем. Покажем, что она не интегрируема. Зафиксируем произвольное разбиение $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ отрезка [0, 1]. Если выбрать точки $\xi_i \in [x_{i-1}, x_i]$, $i = 1, 2, \ldots, i_\tau$, рациональными, то получим

$$\sigma_\tau = \sum_{i=2}^{i_\tau} f(\xi_i)\Delta x_i = \sum_{i=2}^{i_\tau} \Delta x_i = 1,$$

а если взять ξ_i иррациональными, то

$$\sigma_\tau = \sum_{i=1}^{i_\tau} f(\xi_i)\Delta x_i = 0.$$

Это верно для любого разбиения τ, следовательно, интегральные суммы σ_τ заведомо не стремятся ни к какому пределу при $|\tau| \to 0$.

23.4. Верхние и нижние суммы Дарбу. Верхний и нижний интегралы Дарбу

Пусть рассматриваемая функция $f(x)$ определена на отрезке $[a, b]$, $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ — некоторое его разбиение и $\Delta x_i = x_i - x_{i-1}$, $i = 1, 2, \ldots, i_\tau$. Положим (рис. 107)

$$M_i = \sup_{x_{i-1} \leqslant x \leqslant x_i} f(x),\ m_i = \inf_{x_{i-1} \leqslant x \leqslant x_i} f(x),\ i = 1, 2, \ldots, i_\tau,$$

$$S_\tau = S_\tau(f) = \sum_{i=1}^{i_\tau} M_i \Delta x_i,\ s_\tau = s_\tau(f) = \sum_{i=1}^{i_\tau} m_i \Delta x_i. \quad (23.10)$$

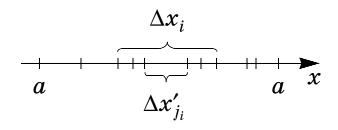


Рис. 108

Очевидно, $s_\tau \leqslant S_\tau$. Сумма S_τ называется *верхней*, а s_τ — *нижней суммой Дарбу функции f*.

При заданной функции $f(x)$ ее верхняя S_τ и нижняя s_τ суммы Дарбу являются функциями разбиений отрезка $[a, b]$. Поэтому для них имеет смысл понятие пределов $\lim\limits_{|\tau| \to 0} S_\tau$ и $\lim\limits_{|\tau| \to 0} s_\tau$ (см. определение 3 в п. 23.1).

Свойства сумм Дарбу

1^0. *Если функция f ограничена, то при любом разбиении суммы S_τ и s_τ определены.*

В самом деле, в этом случае M_i и m_i, $i = 1, 2, \ldots, i_\tau$, конечны, поэтому выражения (23.10) имеют смысл.

2^0. *Если $\tau' \succ \tau$, то $S_{\tau'} \leqslant S_\tau$ и $s_\tau \leqslant s_{\tau'}$.*

Доказательство. Пусть $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ и $\tau' = \{x_j'\}_{j=0}^{j=j_{\tau'}}$ — два разбиения отрезка $[a, b]$ таких, что $\tau \prec \tau'$ и

$$m_i = \inf_{x_{i-1} \leqslant x \leqslant x_i} f(x),\ i = 1, 2, \ldots, i_\tau,$$

$$m_j' = \inf_{x_{j-1}' \leqslant x \leqslant x_j'} f(x),\ j = 1, 2, \ldots, j_j.$$

Если $[x_{j-1}', x_j'] \subset [x_{i-1}, x_i]$, то, очевидно,

$$m_i \leqslant m_j' \qquad (23.11)$$

(нижняя грань при уменьшении множества может только увеличиться).

В силу условия $\tau \prec \tau'$, каждый отрезок $[x_{i-1}, x_i]$ разбиения τ является объединением каких-то отрезков разбиения τ'; будем обозначать эти отрезки через $[x'_{(j-1)_i}, x'_{j_i}]$. Таким образом, если $\Delta x_i = x_i - x_{i-1}$ и $\Delta x'_{j_i} = x'_{j_i} - x'_{(j-1)_i}$, то (рис. 108) $\Delta x_i = \sum\limits_{j_i} \Delta x'_{j_i}$. Используя эти обозначения и неравенство (23.11), получим

$$s_\tau = \sum_{i=1}^{i_\tau} m_i \Delta x_i = \sum_{i=1}^{i_\tau} m_i \sum_{j_i} \Delta x'_{j_i} = \sum_{i=1}^{i_\tau} \sum_{j_i} m_i \Delta x'_{j_i} \leqslant$$
$$\leqslant \sum_{i=1}^{i_\tau} \sum_{j_i} m'_{j_i} \Delta x'_{j_i} = \sum_{j=1}^{i_{\tau'}} m_j' \Delta x_j' = s_{\tau'}.$$

Мы доказали, что $s_\tau \leqslant s'_\tau$. Аналогично доказывается, что $S_\tau \geqslant S'_\tau$ при $\tau \prec \tau'$. □

СЛЕДСТВИЕ. *Для любых двух разбиений* τ_1 *и* τ_2 *отрезка* $[a, b]$ *выполняется неравенство*

$$s_{\tau_1} \leqslant S_{\tau_2}, \qquad (23.12)$$

т. е. любая нижняя сумма Дарбу не больше любой верхней.

Действительно, если даны два разбиения τ_1 и τ_2 отрезка $[a, b]$, то существует разбиение τ этого отрезка такое, что $\tau \succ \tau_1$ и $\tau \succ \tau_2$ (см. п. 23.1). Применяя свойство 2^0, получаем

$$s_{\tau_1} \leqslant s_\tau \leqslant S_\tau \leqslant S_{\tau_2}. \ \square$$

В силу неравенств $m_i \leqslant f(\xi_i) \leqslant M_i$, $x_{i-1} \leqslant \xi_i \leqslant x_i$, $i = 1, 2, \dots, i_r$, суммы Римана и Дарбу связаны неравенствами $s_\tau \leqslant \sigma_\tau \leqslant S_\tau$.

Следующее свойство является уточнением этого утверждения.

3^0. *Если* $\sigma_\tau = \sigma_\tau(f; \xi_1, \dots, \xi_{i_\tau})$ — *какая-либо интегральная сумма Римана, соответствующая данному разбиению* τ, *то*

$$s_\tau = \inf_{\xi_1, \dots, \xi_{i_\tau}} \sigma_\tau, \quad S_\tau = \sup_{\xi_1, \dots, \xi_{i_\tau}} \sigma_\tau.$$

Доказательство. Пусть $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ — разбиение отрезка $[a, b]$ и $\xi_i \in [x_{i-1}, x_i]$, $i = 1, 2, \dots, i_\tau$.

Если заданы какие-либо числовые множества X_i, и постоянные $a_i > 0$, $i = 1, 2, \dots, i_0$, то для множества

$$X = \left\{x : x = \sum_{i=1}^{i_0} a_i x_i,\ x \in X_i,\ i = 1, 2, \dots, i_0\right\}$$

справедливы равенства (см. п. 3.5*)

$$\sup X = \sum_{i=1}^{i_0} a_i \sup X_i,\ \inf X = \sum_{i=1}^{i_0} a_i \inf X_i.$$

В силу этого имеем

$$s_\tau = \sum_{i=1}^{i_\tau} m_i \Delta x_i = \sum_{i=1}^{i_\tau} [\inf_{x_{i-1} \leqslant \xi_i \leqslant x_i} f(\xi_i)] \Delta x_i =$$

$$= \inf_{\substack{x_{i-1} \leqslant \xi_i \leqslant x_i \\ i = 1, 2, \dots, i_\tau}} \sum_{i=1}^{i_\tau} f(\xi_i) \Delta x_i = \inf_{\substack{x_{i-1} \leqslant \xi_i \leqslant x_i \\ i = 1, 2, \dots, i_\tau}} \sigma_\tau(f; \xi_1, \dots, \xi_{i_\tau}).$$

Аналогично,

$$S_\tau = \sum_{i=1}^{i_\tau} M_i \Delta x_i = \sum_{i=1}^{i_\tau} [\sup_{x_{i-1} \leqslant \xi_i \leqslant x_i} f(\xi_i)] \Delta x_i =$$

$$= \sup_{\substack{x_{i-1} \leqslant \xi_i \leqslant x_i \\ i = 1, 2, \ldots, i_\tau}} \sum_{i=1}^{i_\tau} f(\xi_i) \Delta x_i = \sup_{\substack{x_{i-1} \leqslant \xi_i \leqslant x_i \\ i = 1, 2, \ldots, i_\tau}} \sigma_\tau(f; \xi_1, \ldots, \xi_{i_\tau}). \ \square$$

4^0. $S_\tau - s_\tau = \sum_{i=1}^{i_\tau} \omega_i(f) \Delta x_i$, *где* $\omega_i(f)$ — *колебание функции* f *на отрезке* $[x_{i-1}, x_i]$ (см. п. 6.4), $i = 1, 2, \ldots, i_\tau$.

Доказательство. Отметим сначала, что если для двух данных числовых множеств X и Y положить

$$Z = \{z : z = x - y, x \in X, y \in Y\},$$

то $\sup Z = \sup X - \inf Y$ (см. п. 3.5*). Используя это, получим

$$M_i - m_i = \sup_{x_{i=1} \leqslant x \leqslant x_i} f(x) - \inf_{x_{i=1} \leqslant x \leqslant x_i} f(x) =$$

$$= \sup_{\substack{x_{i-1} \leqslant x' \leqslant x_i \\ x_{i-1} \leqslant x'' \leqslant x_i}} [f(x'') - f(x')] = \omega_i(f), \quad i = 1, 2, \ldots, i_\tau,$$

поэтому

$$S_\tau - s_\tau = \sum_{i=1}^{i_\tau} (M_i - m_i) \Delta x_i = \sum_{i=1}^{i_\tau} \omega_i(f) \Delta x_i. \ \square$$

Положим теперь $I_* = \sup_\tau s_\tau$, $I^* = \inf_\tau S_\tau$.

Число или отрицательная бесконечность I_* называется *нижним интегралом Дарбу* функции f на отрезке $[a, b]$, а число или положительная бесконечность I^* — ее *верхним интегралом*.

Из свойств 1^0 и 2^0 сумм Дарбу следует, что если функция ограничена, то как ее нижний интеграл Дарбу, так и верхний конечны. В силу следствия из свойства 2^0 будем также иметь

$$I_* \leqslant I^*. \qquad (23.13)$$

В самом деле, перейдя в левой части неравенства (23.12) к верхней грани по разбиениям τ_1, получим, что для любого разбиения τ_2 выполняется неравенство $I_* \leqslant S_{\tau_2}$. Перейдя в этом неравенстве к нижней грани по разбиениям τ_2, будем иметь неравенство $I_* \leqslant I^*$.

23.5. Необходимые и достаточные условия интегрируемости

ТЕОРЕМА 3. *Для того чтобы ограниченная на некотором отрезке функция была интегрируемой на нем, необходимо и достаточно, чтобы*

$$\lim_{|\tau| \to 0} (S_\tau - s_\tau) = 0. \qquad (23.14)$$

Условие (23.14) означает (см. определение 3 в п. 23.1), что для всякого $\varepsilon > 0$ существует такое $\delta > 0$, что для любого разбиения τ мелкости $|\tau| < \delta$ выполняется неравенство

$$|S_\tau - s_\tau| < \varepsilon. \qquad (23.15)$$

Так как $s_\tau \leqslant S_\tau$, то (23.15) равносильно неравенству

$$S_\tau - s_\tau < \varepsilon.$$

Доказательство необходимости. Пусть ограниченная на отрезке $[a, b]$ функция f интегрируема на нем и пусть $I = \int\limits_a^b f(x)\,dx$; тогда $\lim\limits_{|\tau| \to 0} \sigma_\tau = I$. Поэтому для любого $\varepsilon > 0$ существует такое $\delta > 0$, что если $|\tau| < \delta$, то

$$|\sigma_\tau - I| < \varepsilon \text{ или } I - \varepsilon < \sigma_\tau < I + \varepsilon.$$

Отсюда при $|\tau| < \delta$, согласно свойству 3^0 сумм Дарбу (см. п. 23.4), получаем неравенство

$$I - \varepsilon \leqslant s_\tau \leqslant S_\tau \leqslant I + \varepsilon.$$

Таким образом, если $|\tau| < \delta$, то $0 \leqslant S_\tau - s_\tau \leqslant 2\varepsilon$, а это и означает выполнение условия (23.14).

Доказательство достаточности. Пусть функция f ограничена и выполняется условие (23.14). Из определения нижнего и верхнего интегралов Дарбу и из неравенства (23.13) имеем

$$s_\tau \leqslant I_* \leqslant I^* \leqslant S_\tau, \qquad (23.16)$$

поэтому $0 \leqslant I^* - I_* \leqslant S_\tau - s_\tau$, откуда, в силу (23.14), следует, что $I^* - I_* = 0$. Обозначая общее значение верхнего и нижнего интегралов Дарбу через I, т. е. полагая $I = I_* = I^*$, из (23.16) получаем $s_\tau \leqslant I \leqslant S_\tau$, поэтому

$$0 \leqslant I - s_\tau \leqslant S_\tau - s_\tau, \quad 0 \leqslant S_\tau - I \leqslant S_\tau - s_\tau.$$

Отсюда, в силу (23.14), имеем

$$\lim_{|\tau| \to 0} (I - s_\tau) = \lim_{|\tau| \to 0} (S_\tau - I) = 0,$$

следовательно,

$$\lim_{|\tau| \to 0} s_\tau = \lim_{|\tau| \to 0} S_\tau = I. \tag{23.17}$$

Но, в силу свойства интегральных сумм Дарбу (см. п. 23.4),

$$s_\tau \leqslant \sigma_\tau \leqslant S_\tau. \tag{23.18}$$

Из (23.17) и (23.18) следует (см. п. 23.1), что $\lim\limits_{|\tau| \to 0} \sigma_\tau = I$, а это и означает интегрируемость функции f. □

СЛЕДСТВИЕ 1. *Если функция f интегрируема, то не только ее интегральные суммы Римана, но и ее суммы Дарбу стремятся к ее интегралу при стремлении мелкости разбиения к нулю.*

Действительно, если функция f интегрируема, то выполняется условие (23.14), а из него, как мы видели, и следует утверждение следствия, т. е. равенство (23.17). □

СЛЕДСТВИЕ 2. *Для того чтобы ограниченная на некотором отрезке функция f была интегрируемой на этом отрезке, необходимо и достаточно, чтобы*

$$\lim_{|\tau| \to 0} \sum_{i=1}^{i_\tau} \omega_i(f) \Delta x_i = 0,$$

где $\omega_i(f)$ — колебание функции f на отрезке $[x_{i-1}, x_i]$ разбиения $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ отрезка $[a, b]$.

Это следует непосредственно из теоремы 3 и свойства 4^0 сумм Дарбу (см. п. 23.4).

23.6. Интегрируемость непрерывных и монотонных функций

ТЕОРЕМА 4. *Функция, определенная и непрерывная на некотором отрезке, интегрируема на нем.*

Доказательство. Пусть функция f непрерывна на отрезке $[a, b]$; тогда, как известно, она ограничена (см. теорему 1 в п. 6.1) и равномерно непрерывна (см. теорему 5 в п. 6.4) на этом отрезке. Зафиксируем произвольно $\varepsilon > 0$.

В силу равномерной непрерывности существует такое $\delta > 0$, что для любых точек $\xi \in [a, b]$ и $\eta \in [a, b]$, удовлетворяющих условию $|\eta - \xi| < \delta$, выполняется неравенство

$$|f(\eta) - f(\xi)| < \frac{\varepsilon}{b - a}. \tag{23.19}$$

Возьмем какое-либо разбиение $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ мелкости $|\tau| < \delta$. Пусть, как всегда, $\Delta x_i = x_i - x_{i-1}$, $m_i = \inf\limits_{[x_{i-1}, x_i]} f(x)$, $M_i = \sup\limits_{[x_{i-1}, x_i]} f(x)$, $i = 1, 2, \ldots, i_\tau$. Непрерывная на отрезке функция достигает своей нижней и верхней грани на этом отрезке, поэтому существуют такие точки $\xi_i \in [x_{i-1} - x_i]$ и $\eta_i \in [x_{i-1}, x_i]$, что

$$f(\xi_i) = m_i, \quad f(\eta_i) = M_i.$$

Точки ξ_i и η_i принадлежат одному и тому же отрезку разбиения τ, следовательно,

$$|\eta_i - \xi_i| \leqslant \Delta x_i \leqslant |\tau| < \delta.$$

Отсюда, в силу (23.19), вытекает неравенство

$$f(\eta_i) - f(\xi_i) = |f(\eta_i) - f(\xi_i)| < \frac{\varepsilon}{b - a}, \; i = 1, 2, \ldots, i_\tau.$$

Следовательно, для любого разбиения τ мелкости $|\tau| < \delta$ выполняется условие

$$0 \leqslant S_\tau - s_\tau = \sum_{i=1}^{i_\tau} (M_i - m_i)\Delta x_i =$$
$$= \sum_{i=1}^{i_\tau} [f(\eta_i) - f(\xi_i)]\Delta x_i < \frac{\varepsilon}{b - a} \sum_{i=1}^{i_\tau} \Delta x_i = \varepsilon.$$

Это означает, что $\lim\limits_{|\tau| \to 0} (S_\tau - s_\tau) = 0$. Поэтому, согласно теореме 3, функция f интегрируема на отрезке $[a, b]$. □

ТЕОРЕМА 5. *Функция, определенная и монотонная на отрезке* $[a, b]$, *интегрируема на этом отрезке.*

Доказательство. Пусть функция $f(x)$ монотонна на отрезке $[a, b]$, например возрастает на нем. Тогда

$$f(a) \leqslant f(x) \leqslant f(b), \quad a \leqslant x \leqslant b.$$

Таким образом, функция f ограничена на отрезке $[a, b]$. Далее, для любого разбиения $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ отрезка $[a, b]$, очевидно, имеем $m_i = f(x_{i-1})$, $M_i = f(x_i)$, $i = 1, 2, \ldots, i_\tau$, поэтому

$$S_\tau(f) - s_\tau(f) = \sum_{i=1}^{i_\tau} (M_i - m_i)\Delta x_i = \sum_{i=1}^{i_\tau} [f(x_i\} - f(x_{i-1})]\Delta x_i \leqslant$$
$$\leqslant |\tau| \sum_{i=1}^{i_\tau} [f(x_i) - f(x_{i-1})] = [f(b) - f(a)]|\tau|,$$

так как в сумме $\sum_{i=1}^{i_\tau}[f(x_i) - f(x_{i-1})]$ взаимно уничтожаются все слагаемые, кроме $f(b)$ и $f(a)$. Из полученного неравенства следует, что

$$\lim_{|\tau| \to 0} [S_\tau(f) - s_\tau(f)] = 0.$$

Поэтому (см. п. 23.5) функция f интегрируема на отрезке $[a, b]$. □

Монотонные функции могут иметь счетное множество точек разрыва и вместе с тем, согласно доказанной теореме, быть интегрируемыми.

Примером немонотонной функции, имеющей счетное множество точек разрыва и вместе с тем интегрируемой, является функция Римана, которая была определена в п. 5.12. Докажем это.

Напомним, что упомянутая функция f равна $1/n$ в каждой, не равной нулю, рациональной точке $r = m/n$ (где m/n — несократимая рациональная дробь) отрезка [0, 1] и равна нулю во всех остальных точках этого отрезка. Покажем, что

$$\int_0^1 f(x)\,dx = 0.$$

Зададим произвольно $\varepsilon > 0$. Выберем натуральное число n так, чтобы

$$\frac{1}{n} < \frac{\varepsilon}{2}, \tag{23.20}$$

и число $\delta > 0$ так, чтобы

$$\delta < \frac{1}{n^3}. \tag{23.21}$$

Для любого разбиения $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ мелкости $|\tau| \leqslant \delta$ его отрезок назовем отрезком первого рода, если он не содержит чисел вида

$$\frac{m}{p}, \quad m = 1, 2, \ldots, p \leqslant n, \tag{23.22}$$

и второго рода, если он содержит хотя бы одну из указанных точек. Ясно, что каждый отрезок разбиения τ является либо отрезком первого рода, либо второго.

Всякую интегральную сумму σ_τ функции f представим в виде суммы двух слагаемых следующим образом:

$$\sigma_\tau = \sum_{i=1}^{i_\tau} f(\xi_i)\Delta x_i = \Sigma' f(\xi_i)\Delta x_i + \Sigma'' f(\xi_i)\Delta x_i, \qquad (23.23)$$

где знак штрих у сигмы означает, что суммирование производится только по отрезкам первого рода, а два штриха у сигмы — что суммирование производится только по отрезкам второго рода. Оценим отдельно каждое слагаемое правой части равенства (23.23).

В точках отрезков первого рода, согласно заданию функции f, все ее значения меньше $1/n$, следовательно,

$$|\Sigma' f(\xi_i)\Delta x_i| \leqslant \Sigma' |f(\xi_i)|\Delta x_i < \frac{1}{n}\Sigma' \Delta x_i \leqslant$$

$$\leqslant \frac{1}{n}\sum_{i=1}^{i_\tau} \Delta x_i = \frac{1}{n} \underset{(23.20)}{<} \frac{\varepsilon}{2}. \qquad (23.24)$$

Для оценки второго слагаемого заметим, что для всех точек x отрезка $[0, 1]$, согласно определению функции f, имеет место неравенство $|f(x)| \leqslant 1$ и что число отрезков второго рода не превышает n^2, так как число всех точек (23.22), очевидно, не превышает суммы $1 + 2 + \ldots + n = \frac{n(n-1)}{2} \leqslant \frac{n^2}{2}$, а каждая точка может принадлежать не более чем двум отрезкам разбиения τ. Поэтому

$$|\Sigma'' f(\xi_i)\Delta x_i| \leqslant \Sigma'' |f(\xi_i)|\Delta x_i \leqslant \Sigma'' \Delta x_i \leqslant n^2|\tau| \leqslant$$

$$\leqslant n^2\delta \underset{(23.21)}{<} \frac{1}{n} \underset{(23.20)}{<} \frac{\varepsilon}{2}. \qquad (23.25)$$

Из соотношений (23.23), (23.24) и (23.25) следует, что для любой интегральной суммы σ_τ, соответствующей разбиению τ мелкости $|\tau| < \delta$, выполняется неравенство $|\sigma_\tau| < \varepsilon$, что, в силу произвольности $\varepsilon > 0$, и означает, что $\int_0^1 f(x)\,dx = 0$.

23.7*. Критерии интегрируемости Дарбу и Римана

Покажем, что для ограниченных на отрезке функций нижний I_* и верхний I^* интегралы Дарбу (см. п. 23.4) являются

не только верхней и нижней гранями соответственно нижних и верхних сумм Дарбу, но и их пределами.

ТЕОРЕМА 6. *Если функция f ограничена на отрезке, то*

$$I_* = \lim_{|\tau| \to 0} s_\tau, \quad I^* = \lim_{|\tau| \to 0} S_\tau. \tag{23.26}$$

Здесь, как обычно, s_τ и S_τ — соответственно нижняя и верхняя суммы Дарбу функции f, соответствующие разбиению τ отрезка $[a, b]$.

Доказательство. Докажем справедливость первой из формул (23.26) (вторая доказывается аналогично). В силу ограниченности функции f на отрезке $[a, b]$, существует такая постоянная $c > 0$, что

$$|f(x)| \leqslant c, \quad x \in [a, b]. \tag{23.27}$$

Зададим произвольно $\varepsilon > 0$. Согласно определению нижнего интеграла (см. п. 23.4), т. е., согласно равенству $I_* = \sup\limits_{\tau} s_\tau$, существует такое разбиение $\tau_0 = [x_i^{(0)}]_{i=0}^{i=i_0}$ отрезка $[a, b]$, что

$$s_{\tau_0} > I_* - \frac{\varepsilon}{2}. \tag{23.28}$$

Положим

$$\delta = \frac{\varepsilon}{4 i_0 c}. \tag{23.29}$$

Пусть $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ — произвольное разбиение отрезка $[a, b]$ мелкости, меньшей δ, т. е. $|\tau| < \delta$. Рассмотрим разбиение $\tau^* = \{x_i^*\}_{i=0}^{i=i^*}$, получающееся объединением разбиений τ_0 и τ. Если $s_\tau = \sum\limits_{i=1}^{i_\tau} m_i \Delta x_i$ и $s_\tau^* = \sum\limits_{i=1}^{i_\tau} m_i^* \Delta x_i^*$ — соответствующие разбиениям τ и τ^* нижние суммы Дарбу, то, поскольку разбиение τ^* вписано в разбиение τ_0, выполняется (см. п. 23.4) неравенство $s_{\tau_0} \leqslant s_\tau^*$, поэтому (см. (23.28)) $s_\tau^* > I_* - \frac{\varepsilon}{2}$, т. е.

$$I_* - s_\tau^* < \frac{\varepsilon}{2}. \tag{23.30}$$

Оценим разность $s_{\tau^*} - s_\tau$. Разбиение τ^* получается из разбиения τ добавлением к нему точек разбиения τ_0. Поэтому тем отрезкам разбиения τ, внутрь которых не попали точки разбиения τ_0, соответствуют одинаковые слагаемые в суммах s_τ

и s^*. Если же внутрь отрезка $[x_{i-1}, x_i]$ разбиения τ попали одна или несколько точек разбиения τ_0, то такой отрезок является объединением отрезков $[x^*_{j_i-1}, x^*_{j_i}]$ разбиения τ^*:

$$\Delta x_i = \sum_{i=1}^{i_\tau} \Delta x^*_{j_i}.$$

Так как

$$|m_i| \underset{(23.27)}{\leqslant} c,\ i = 1, 2, \ldots, i_\tau,$$

$$|m^*_i| \underset{(23.27)}{\leqslant} c,\ i = 1, 2, \ldots, i_{\tau^*}, \qquad (23.31)$$

то

$$\sum_{j_i} m^*_{j_i}\Delta x^*_{j_i} - m_i\Delta x_i =$$

$$= \sum_{j_i} (m^*_{j_i} - m_i)\Delta x^*_{j_i} \underset{(23.31)}{\leqslant} 2c \sum_{j_i} \Delta x^*_{j_i} = 2c\Delta x_i \leqslant 2c\delta. \qquad (23.32)$$

Число таких отрезков не превышает числа i_0 точек разбиения τ_0. Поэтому

$$s^*_\tau - s_\tau \leqslant 2c\delta_{i_0} \underset{(23.29)}{=} \frac{\varepsilon}{2}. \qquad (23.33)$$

Таким образом, если $|\tau| < \delta$, то

$$0 \leqslant I_* - s_\tau = (I_* - s^*_\tau) + (s^*_\tau - s_\tau) \underset{\substack{(23.30)\\(23.33)}}{<} \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Это и означает, что $\lim\limits_{|\tau| \to 0} s_\tau = I_*$. □

ТЕОРЕМА 7 (теорема Дарбу). *Для того чтобы ограниченная на отрезке функция была на нем интегрируема по Риману, необходимо и достаточно, чтобы были равны ее верхний и нижний интегралы Дарбу.*

Доказательство необходимости. Пусть функция f интегрируема на отрезке $[a, b]$. Так как $s_\tau \leqslant I_* \leqslant I^* \leqslant S_\tau$, то $0 \leqslant I^* - I_* \leqslant S_\tau - s_\tau$ и так как для интегрируемой функции (см. (23.14)) $\lim\limits_{|\tau| \to 0} (S_\tau - s_\tau) = 0$, то $I_* = I^*$.

Доказательство достаточности. Если f — ограниченная на отрезке $[a, b]$ функция и $I_* = I^*$, то, в силу теоремы 6,

$$\lim_{|\tau| \to 0} (S_\tau - s_\tau) = \lim_{|\tau| \to 0} S_\tau - \lim_{|\tau| \to 0} s_\tau = I^* - I_* = 0.$$

Отсюда, согласно теореме 3 (см. п. 23.5), функция f интегрируема на отрезке $[a, b]$. □

ТЕОРЕМА 8 (критерий Римана). *Для того чтобы ограниченная на отрезке функция была на нем интегрируема по Риману, необходимо и достаточно, чтобы для любого $\varepsilon > 0$ существовало такое разбиение τ отрезка $[a, b]$, что*

$$S_\tau - s_\tau < \varepsilon. \qquad (23.34)$$

СЛЕДСТВИЕ. *Для того чтобы ограниченная на отрезке $[a, b]$ функция f была интегрируема на нем, необходимо и достаточно, чтобы для любого $\varepsilon > 0$ существовало такое разбиение $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ отрезка $[a, b]$, что*

$$\sum_{i=1}^{i_\tau} \omega_i(f)\Delta x_i < \varepsilon.$$

Здесь $\omega_i(f)$ — колебание функции f на отрезке $[x_{i-1}, x_i]$, $\Delta x_i = x_i - x_{i-1}$, $i = 1, 2, \dots, i_\tau$.

Доказательство необходимости. Если функция f интегрируема на отрезке, то, согласно теоремам 6 и 7, $\lim\limits_{|\tau| \to 0} s_\tau = I_* = I^* = \lim\limits_{|\tau| \to 0} S_\tau$, откуда $\lim\limits_{|\tau| \to 0} (S_\tau - s_\tau) = 0$. Следовательно, для любого $\varepsilon > 0$ найдется такое разбиение τ отрезка $[a, b]$, что $S_\tau - s_\tau < \varepsilon$.

Доказательство достаточности. Если выполняется условие (23.34), то, в силу неравенства $s_\tau \leqslant I_* \leqslant I^* \leqslant S_\tau$, при любом $\varepsilon > 0$ справедливо неравенство $0 \leqslant I^* - I_* \leqslant S_\tau - s_\tau < \varepsilon$. Следовательно, $I_* = I^*$, откуда, согласно теореме 6, вытекает интегрируемость функции f. □

Следствие сразу следует из доказанной теоремы, так как (см. п. 23.4)

$$S_\tau - s_\tau = \sum_{i=1}^{i_\tau} \omega_i(f)\Delta x_i.$$

Теорема 7 называется *критерием Дарбу*, а теорема 8 — *критерием Римана* интегрируемости функции.

Замечание 1. Критерий Римана удобнее использовать для выяснения интегрируемости функции, чем теорему 3 (см. п. 23.5), так как здесь достаточно найти хотя бы одно разбиение τ отрезка, удовлетворяющее условию (23.34),

тогда как и в первом случае надо проверять это условие для всех достаточно мелких разбиений.

С помощью критерия Римана легко, например, устанавливается, что если функция задана на отрезке $[a, b]$ и интегрируема на отрезках $[a, c]$ и $[c, b]$, $a < c < b$, то она интегрируема и на отрезке $[a, b]$.

В самом деле, прежде всего из интегрируемости функции f на отрезках $[a, c]$ и $[c, b]$ следует ее ограниченность на этих отрезках, а потому и ограниченность на всем отрезке $[a, b]$. Далее, из интегрируемости функции f на отрезках $[a, c]$ и $[c, b]$ следует, в силу критерия Римана, что для любого $\varepsilon > 0$ существуют такие разбиения τ_1 и τ_2 соответственно отрезков $[a, c]$ и $[c, b]$, что соответствующие им верхние и нижние суммы Дарбу отличаются друг от друга меньше, чем на $\varepsilon/2$, но тогда верхняя и нижняя суммы Дарбу, соответствующие разбиению отрезка $[a, b]$, которые получаются объединением разбиений τ_1 и τ_2, отличаются меньше чем на ε. Согласно критерию Римана, это означает, что рассматриваемая функция интегрируема на отрезке $[a, b]$.

Замечание 2. В качестве другого примера применения критерия интегрируемости Римана докажем, что функция f, ограниченная на некотором отрезке $[a, b]$ и интегрируемая по Риману на любом отрезке $[a, \eta]$, $a < \eta < b$, интегрируема по Риману и на всем отрезке $[a, b]$. Действительно, если $|f(x)| \leqslant M$, $x \in [a, b]$, и задано $\varepsilon > 0$, то выберем δ, $0 < \delta < b - a$, так, чтобы $\delta < \frac{\varepsilon}{2M}$. Тогда, в силу интегрируемости функции f на отрезке $[a, b - \delta]$, существует такое его разбиение τ, что если s_τ и S_τ — нижняя и верхняя суммы Дарбу функции для этого разбиения, то $S_\tau - s_\tau < \varepsilon/2$.

Обозначим через τ разбиение отрезка $[a, b]$, получающееся из разбиения τ_0 отрезка $[a, b - \delta]$ добавлением точки b: $\tau_0 = \tau \cup \{b\}$, и пусть $m_0 = \inf\limits_{[b-\delta,\, b]} f(x)$, $M_0 = \sup\limits_{[b-\delta,\, b]} f(x)$. Если s_{τ_0} и S_{τ_0} — нижняя и верхняя суммы Дарбу функции для разбиения τ_0, то

$$S_{\tau_0} = S_\tau + M_0\delta, \ \ s_{\tau_0} = S_\tau + m_0\delta.$$

Поэтому

$$S_{\tau_0} - s_{\tau_0} = S_\tau - s_\tau + (M_0 - m_0)\delta < \frac{\varepsilon}{2} + M\delta = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

и, следовательно, согласно теореме 8, функция f интегрируема по Риману на отрезке $[a, b]$.

23.8*. Колебания функций

По аналогии с определением колебания функции на отрезке (см. п. 6.4) определяется колебание функции на множестве. Пусть функция f определена на множестве $X \subset \boldsymbol{R}$. Ее колебание $\omega(f; X)$ на множестве X определяется равенством

$$\omega(f; X) = \sup_{x,\, x' \in X} |f(x') - f(x)|. \qquad (23.35)$$

Очевидно, что если $X_1 \subset X$, то

$$\omega(f; X_1) \leqslant \omega(f; X) \qquad (23.36)$$

(при переходе от множества к его подмножеству верхняя грань не возрастает), и что для любых двух пересекающихся подмножеств X', X'' множества $X : X' \subset X,\ \ X'' \subset X$, $X' \cap X'' \neq \varnothing$ выполняется неравенство

$$\omega(f; X' \cup X'') \leqslant \omega(f; X') + \omega(f; X''). \qquad (23.37)$$

В дальнейшем под окрестностью точки будем понимать любой интервал числовой оси, содержащей эту точку (см. п. 3.2).

Определение 4. *Колебанием* $\omega(f; x_0)$ *функции* f *в точке* $x_0 \in \in X$ *называется нижняя грань колебаний функции на пересечениях всевозможных окрестностей* $U(x_0)$ *этой точки с множеством* X:

$$\omega(f; x_0) \overset{\text{def}}{=} \inf_{U(x_0)} \omega(f; U(x_0) \cap X). \qquad (23.38)$$

Очевидно, что

$$0 \leqslant \omega(f; x_0) \leqslant +\infty.$$

ЛЕММА 1. *Если функция* f *ограничена на множестве* X, *т. е. если* $|f(x)| \leqslant c$, *то для всех* $x \in X$ *выполняется неравенство*

$$\omega(f; x) \leqslant 2c. \qquad (23.39)$$

Доказательство. В самом деле,

$$\omega(f; x_0) \underset{(23.38)}{=} \inf_{U(x_0)} \omega(f; U(x_0) \cap X) =$$

$$= \inf_{U(x_0)} \sup_{x,\, x' \in U(x_0) \cap X} |f(x') - f(x)| \leqslant$$

$$\leqslant \inf_{U(x_0)} \sup_{x,\, x' \in U(x_0) \cap X} (|f(x')| + |f(x)| \leqslant 2c. \ \square$$

Для дальнейшего полезно ввести множество

$$X_\varepsilon \overset{\text{def}}{=} \{x \in X : \omega(f;\, x) \geqslant \varepsilon\}, \tag{23.40}$$

где $\varepsilon > 0$ произвольно.

Если $\eta < \varepsilon$, то ясно, что из неравенства $\omega(f;\, x) \geqslant \varepsilon$ следует неравенство $\omega(f;\, x) \geqslant \eta$, поэтому

$$X_\varepsilon \subset X_\eta. \tag{23.41}$$

ЛЕММА 2. *Функция f непрерывна в точке $x \in X$ тогда и только тогда, когда*

$$\omega(f;\, x) = 0. \tag{23.42}$$

СЛЕДСТВИЕ. *Если X_0 — множество точек разрыва функции f, то*

$$X_0 = \bigcup_{n=1}^{\infty} X_{1/n}. \tag{23.43}$$

Доказательство леммы. Если функция f непрерывна в точке $x_0 \in X$, то для любого $\varepsilon > 0$ существует так, что для всех точек $x \in U_\varepsilon(x_0) \cap X$, выполняется неравенство $|f(x) - f(x_0)| < \frac{\varepsilon}{2}$. Поэтому для любых точек $x, x' \in U_\varepsilon(x_0) \cap X$ имеем

$$|f(x') - f(x)| \leqslant \\ \leqslant |f(x') - f(x_0)| + |f(x_0) - f(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \tag{23.44}$$

и, следовательно,

$$\omega(f;\, x_0) = \inf_{U(x_0)} \omega(f;\, U(x_0) \cap X) \leqslant \omega(f;\, \cap X) = \\ = \sup_{x,\, x' \in U_\varepsilon(x_0) \cap X} |f(x') - f(x)| \underset{(23.44)}{\leqslant} \varepsilon.$$

А так как $\varepsilon > 0$ произвольно, то это означает, что $\omega(f;\, x_0) = 0$.

Наоборот, если $\omega(f;\, x_0) = 0$, то для любого $\varepsilon > 0$ существует такая окрестность $U(x_0)$ точки x_0, что $\omega(f;\, U(x_0) \cap X) < \varepsilon$. Тогда для любого $x \in U_\varepsilon(x_0) \cap X$ будем иметь

$$|f(x) - f(x_0)| \underset{(23.35)}{\leqslant} \omega(f;\, U(x_0) \cap X) < \varepsilon,$$

т. е. функция f непрерывна в точке x_0. □

Докажем следствие. Если точка $x_0 \in X$ является точкой разрыва функции f, то, в силу леммы 2, $\omega(f;\, x_0) > 0$, а поэто-

му $x_0 \in X_\varepsilon$ при $\varepsilon = \omega(f;\, x_0)$. Отсюда следует, что множество X_0 точек разрыва функции f представимо в виде

$$X_0 = \bigcup_{\varepsilon > 0} X_\varepsilon. \tag{23.45}$$

Ясно, что $\bigcup\limits_{n=1}^{\infty} X_{1/n} \subset \bigcup\limits_{\varepsilon > 0} X_\varepsilon$, ибо каждое слагаемое левой части включения является и слагаемым правой. С другой стороны, если для данного $\varepsilon > 0$ выбрать так натуральное n, чтобы $1/n < \varepsilon$, то, в силу включения (23.41), будем иметь $X_\varepsilon \subset X_{1/n}$ и, следовательно, $\bigcup\limits_{\varepsilon > 0} X_\varepsilon \subset \bigcup\limits_{n} X_{1/n}$. Таким образом,

$$\bigcup_{n=1}^{\infty} X_{1/n} = \bigcup_{\varepsilon > 0} X_\varepsilon \underset{(23.45)}{=} X_0. \ \square$$

Определение 5. *Множество, состоящее из всех точек прикосновения данного множества, называется его замыканием.*

Для множества X его замыкание обозначается $\overline{X}$.

Поскольку всякая точка множества X является его точкой прикосновения (любая окрестность точки содержит саму точку, т. е. в данном случае содержит точку множества X), то $X \subset \overline{X}$.

Определение 6. *Множество, совпадающее со своим замыканием, называется замкнутым.*

Таким образом, если множество X замкнуто, то $X = \overline{X}$.

ЛЕММА 3. *При любом $\varepsilon > 0$ все точки прикосновения множества X_ε, содержащиеся в X, содержатся и в X_ε, т. е. если $x \in \overline{X}_\varepsilon \cap X$, то $x \in X_\varepsilon$.*

СЛЕДСТВИЕ 1. *Если множество X, на котором задана функция f, замкнутое, то при любом $\varepsilon > 0$ множество X_ε также замкнутое.*

СЛЕДСТВИЕ 2. *Если функция f задана на отрезке $X = [a, b]$, то при любом $\varepsilon > 0$ множество X_ε является ограниченным замкнутым множеством.*

Доказательство леммы. Пусть $\varepsilon > 0$ и $x_0 \in \overline{X}_\varepsilon \cap X$. Зададим произвольно $\eta > 0$. В силу определения (23.38) коле-

бания функции в точке, существует такая окрестность $U(x_0)$ точки x_0, что

$$\omega(f;\, U(x_0) \cap X) < \omega(f;\, x_0) + \eta. \qquad (23.46)$$

Точка x_0 является точкой прикосновения множества X_ε, поэтому существует такая последовательность $x_n \in X_\varepsilon$, $n = 1, 2, \ldots$, что $\lim\limits_{n \to \infty} x_n = x_0$, следовательно, найдется такой номер n_0, что $x_{n_0} \in U(x_0) \cap X_\varepsilon$. Согласно определению колебания функции в точке, отсюда вытекает, что

$$\omega(f;\, U(x_0) \cap X) \geqslant \omega(f;\, x_{n_0}) \qquad (23.47)$$

(окрестность $U(x_0)$ является и окрестностью точки x_{n_0}).

Таким образом,

$$\omega(f;\, x_0) \underset{(23.46)}{>} \omega(f;\, U(x_{n_0}) \cap X) - \eta \underset{(23.47)}{\geqslant}$$

$$\underset{(23.47)}{\geqslant} \omega(f;\, x_{n_0}) - \eta \geqslant \varepsilon - \eta, \qquad (23.48)$$

ибо $x_{n_0} \in X_\varepsilon$, следовательно, $\omega(f;\, x_{n_0}) \geqslant \varepsilon$.

Так как $\omega(f;\, x_0) \underset{(23.48)}{>} \varepsilon - \eta$ при любом $\eta > 0$, то $\omega(f;\, x_0) \geqslant \varepsilon$, т. е. $x_0 \in X_\varepsilon$. Лемма доказана.□

Следствие 1 вытекает из того, что всякая точка прикосновения подмножества (в данном случае X_ε) является и точкой прикосновения самого множества (в данном случае X), поэтому если $x \in \overline{X}_\varepsilon$, то $x \in \overline{X}$. В случае замкнутого множества X имеем $\overline{X} = X$, поэтому $x \in \overline{X}_\varepsilon \cap X$, а тогда, согласно лемме 3, $x \in X_\varepsilon$, т. е. множество X_ε содержит все свои точки прикосновения, что и означает его замкнутость.

Следствие 2 вытекает из того, что отрезок является ограниченным множеством и, следовательно, любое его подмножество также ограничено. □

Прежде чем изучать дальнейшие свойства колебаний функций, докажем одну геометрическую лемму. Предварительно введем понятие покрытия.

Система (конечная или бесконечная) интервалов такая, что каждая точка некоторого заданного множества принадлежит по крайней мере одному из интервалов системы, называется *покрытием* этого множества интервалами.

Таким образом, то, что система $\Omega = \{(a_\alpha, b_\alpha)\}$ интервалов (a_α, b_α), $\alpha \in \mathfrak{A}$, где $\mathfrak{A}$ — некоторое множество индексов, является покрытием множества X, означает, что:

$$X \subset \bigcup_{\alpha \in 0} (a_\alpha, b_\alpha).$$

ЛЕММА 4. *Из всякого покрытия ограниченного замкнутого множества интервалами можно выделить его конечное покрытие.*

Эта лемма обычно называется *леммой Гейне—Бореля*[1]. В дальнейшем в п. 35.3 будет дано ее обобщение на случай покрытий множеств существенно более общего вида.

Доказательство. Пусть Ω — покрытие интервалами ограниченного замкнутого множества X. В силу ограниченности множества X существует отрезок $[a, b]$, в котором содержится это множество: $X \subset [a, b]$. Допустим, что из покрытия Ω нельзя выделить конечного покрытия множества X. Разделим отрезок $[a, b]$ точкой $(a + b)/2$ на два отрезка. Очевидно, покрытие Ω является покрытием пересечения множества X как с одним, так и с другим отрезком. Причем из покрытия Ω нельзя выделить конечного покрытия по крайней мере для одного из этих пересечений. Обозначим такой отрезок $[a_1, b_1]$. Разделим его на два отрезка точкой $\frac{(a_1 + b_1)}{2}$. Продолжая этот процесс, получаем последовательность вложенных отрезков $[a_n, b_n]$ длин $b_n - a_n = \frac{(b - a)}{2^n}$, стремящихся к нулю. Для каждого отрезка $[a_n, b_n]$, $n = 1, 2, \ldots$, система Ω является покрытием пересечения $[a_n, b_n] \cap X$ интервалами, из которого нельзя выделить конечного покрытия этого пересечения.

Если точка ξ является общей точкой, принадлежащей всем отрезкам $[a_n, b_n]$, $n = 1, 2, \ldots$, то она принадлежит множеству X. Действительно, пересечение каждого отрезка $[a_n, b_n]$ с множеством X не пусто. Поэтому существуют точки $x_n \in X$, $a_n \leqslant x_n \leqslant b_n$, $n = 1, 2, \ldots$. Поскольку $\lim\limits_{n \to \infty} a_n = \lim\limits_{n \to \infty} b_n =$

[1] Э. Борель (1871—1956) — французский математик.

$= \xi$, то и $\lim\limits_{n \to \infty} x_n = \xi$. Таким образом, точка ξ является точкой прикосновения множества X, а так как X — замкнутое множество, то $\xi \in X$. В силу этого точка ξ принадлежит и некоторому интервалу (a_α, b_α) покрытия Ω: $(a_\alpha, b_\alpha) \in \Omega$.

Рис. 109

Для числа $\varepsilon = \min\{\xi - a_\alpha; b_\alpha - \xi\}$ существует такой номер n_0, что $b_{n_0} - a_{n_0} < \varepsilon$. Для отрезка $[a_{n_0}, b_{n_0}]$, очевидно, имеют место включения (рис. 109) $\xi \in [a_{n_0}, b_{n_0}] \cap X \subset (a_\alpha, b_\alpha)$, тем самым пересечение $[a_{n_0}, b_{n_0}] \cap X$ оказалось покрытым интервалом (a_α, b_α) из покрытия Ω.

Иначе говоря, из покрытия Ω пересечения отрезка $[a_{n_0}, b_{n_0}]$ с множеством X можно выделить конечное покрытие (состоящее даже из одного интервала) этого пересечения. Но это противоречит выбору отрезков $[a_n, b_n]$, $n = 1, 2, \ldots$. Полученное противоречие доказывает лемму. □

Установим теперь важное для дальнейшего свойство колебаний функций, заданных на отрезке.

ЛЕММА 5. *Пусть функция f задана на отрезке $[a, b]$ и существует такое $\varepsilon > 0$, что для всех точек x отрезка $[a, b]$ выполняется неравенство*

$$\omega(f; x) < \varepsilon, \tag{23.49}$$

тогда существует такое разбиение $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ отрезка $[a, b]$, что для всех $i = 1, 2, \ldots, i_\tau$ имеет место неравенство

$$\omega_i(f) \equiv \omega(f; [x_{i-1}, x_i]) < \varepsilon. \tag{23.50}$$

СЛЕДСТВИЕ. *В условиях леммы*

$$\sum_{i=1}^{i_\tau} \omega_i(f) \Delta x_i < \varepsilon(b - a). \tag{23.51}$$

Доказательство. В силу выполнения условия (23.49), для любой точки $\xi \in [a, b]$ существует такой интервал $(\xi - \delta_\xi, \xi + \delta_\xi)$, что

$$\omega(f; (\xi - \delta_\xi, \xi + \delta_\xi) \cap [a, b]) < \varepsilon. \tag{23.52}$$

Система интервалов

$$\left(\xi - \frac{1}{2}\delta_\xi, \xi + \frac{1}{2}\delta_\xi\right), \xi \in [a, b], \quad \textbf{(23.53)}$$

образует покрытие отрезка $[a, b]$, и если

$$\Delta_\xi \overset{\text{def}}{=} \left[\xi - \frac{1}{2}\delta_\xi, \xi + \frac{1}{2}\delta_\xi\right] \cap [a, b], \quad \textbf{(23.54)}$$

то

$$\omega(f; \Delta_\xi) \underset{(23.36)}{\leqslant} \omega(f; (\xi - \delta_\xi, \xi + \delta_\xi) \cap [a, b]) \underset{(23.52)}{<} \varepsilon. \quad \textbf{(23.55)}$$

Выделим, согласно лемме Гейне—Бореля, из покрытия (23.53) конечное покрытие

$$\left(\xi_1 - \frac{1}{2}\delta_{\xi_1}, \xi_1 + \frac{1}{2}\delta_{\xi_1}\right), \ldots, \left(\xi_n - \frac{1}{2}\delta_{\xi_n}, \xi_n + \frac{1}{2}\delta_{\xi_n}\right)$$

и обозначим концы промежутков

$$\left(\xi_j - \frac{1}{2}\delta_{\xi_j}, \xi_j + \frac{1}{2}\delta_{\xi_j}\right) \cap [a, b]$$

через α_j и β_j, $j = 1, 2, \ldots, n$.

Пусть $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ — разбиение отрезка $[a, b]$, состоящее из всех точек α_j, β_j. Каждый отрезок $[x_{i-1}, x_i]$ этого разбиения имеет один из следующих видов: $[\alpha_j, \beta_j]$, $[\alpha_j, \alpha_k]$, $[\beta_j, \alpha_k]$, $[\beta_j, \beta_k]$, $j, k = 1, 2, \ldots, n$, и целиком содержится в одном из отрезков $\Delta_{\xi_1}, \ldots, \Delta_{\xi_n}$ (см. (23.54)). Иначе говоря, для каждого отрезка $[x_{i-1}, x_i]$ существует такой отрезок Δ_{j_i}, $1 \leqslant j_i \leqslant n$, что $[x_{i-1}, x_i] \subset \Delta_{j_i}$, поэтому

$$\omega_i(f) \equiv \omega(f; [x_{i-1}, x_i]) \underset{(23.36)}{\leqslant} \omega(f; \Delta_{j_i}) \underset{(23.55)}{<} \varepsilon.$$

Лемма доказана. Из нее сразу следует неравенство (23.51):

$$\sum_{i=1}^{i_\tau} \omega_i(f)\Delta x_i \underset{(23.50)}{\leqslant} \varepsilon \sum_{i=1}^{i_\tau} \Delta x_i = \varepsilon(b - a). \ \square$$

З а м е ч а н и е. Если во всех точках отрезка $[a, b]$ при некотором $\varepsilon > 0$ выполняется неравенство (23.49), то функция f ограничена на этом отрезке. В самом деле, если бы она была на нем не ограничена, то на отрезке нашлась бы такая точка x_0, что в любой ее окрестности функция f была бы не ограничена, а тогда $\omega(f; x_0) = +\infty$.

23.9*. Критерий интегрируемости Дюбуа-Реймона

Докажем теперь теорему, которая называется критерием интегрируемости *Дюбуа-Реймона*[1]. Предварительно сделаем следующее замечание.

Если какое-либо множество числовой прямой нельзя покрыть конечной системой интервалов с суммой их длин меньше некоторого δ, то это множество нельзя покрыть и конечной системой отрезков с суммой их длин меньше $\delta - \eta$, где $\eta > 0$ — произвольное число.

В самом деле, если бы нашлась такая система, состоящая из n отрезков, покрывающих рассматриваемое множество, то, заменив каждый из них содержащим его несколько большим интервалом, а именно таким, чтобы его длина отличалась от длины отрезка не более чем на η/n, получим конечную систему интервалов, покрывающих данное множество с суммой длин меньше δ, что противоречит сделанному предположению.

ТЕОРЕМА 9 (теорема Дюбуа-Реймона). *Ограниченная на отрезке* $[a, b]$ *функция* f *интегрируема на этом отрезке тогда и только тогда, когда для любого* $\varepsilon > 0$ *и любого* $\delta > 0$ *множество всех точек* $x \in [a, b]$, *в которых* $\omega(f; x) \geqslant \varepsilon$, *можно покрыть конечной системой интервалов с суммой длин, меньшей* δ.

Доказательство необходимости. Пусть функция f интегрируема на отрезке $X = [a, b]$ и существуют такие $\varepsilon > 0$ и $\delta > 0$, что, какую бы конечную систему интервалов, покрывающих множество $X_\varepsilon = [a, b]_\varepsilon$ (см. (23.40)), ни взять, сумма их длин всегда не меньше δ.

Пусть $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ — произвольное разбиение отрезка $[a, b]$. Каждая точка x, в которой

$$\omega(f; x) \geqslant \varepsilon, \tag{23.56}$$

как и всякая точка отрезка $[a, b]$, принадлежит одному или двум отрезкам разбиения τ. Если она принадлежит одному

[1] П. Дюбуа-Реймон (1831—1889) — немецкий математик.

отрезку $[x_{i-1}, x_i]$ этого разбиения, то возможны следующие случаи. Точка x является внутренней точкой указанного отрезка, т. е. $x \in (x_{i-1}, x_i)$; тогда

$$\omega_i f \equiv \omega_i(f; [x_{i-1}, x_i]) \geqslant \underset{(23.36)}{\geqslant} \omega(f; (x_{i-1}, x_i)) \underset{(23.38)}{\geqslant} \omega(f; x) \geqslant \varepsilon. \qquad (23.57)$$

Возможны еще два случая, когда точка x принадлежит одному отрезку разбиения τ: либо $x = a$, либо $x = b$. И в том и в другом случае аналогичным образом доказывается неравенство (23.57).

Если точка x принадлежит двум отрезкам разбиения τ, то это возможно только тогда, когда она является их общим концом: $x = x_i \in [x_{i-1}, x_i] \cap [x_i, x_{i+1}]$. В этом случае выполняется по крайней мере одно из неравенств, либо $\omega_i(f) \geqslant \frac{\varepsilon}{2}$, либо $\omega_{i+1}(f) \geqslant \frac{\varepsilon}{2}$, так как если бы $\omega_i(f) < \frac{\varepsilon}{2}$ и $\omega_{i+1}(f) < \frac{\varepsilon}{2}$, то

$$\omega(f; x_i) \underset{(23.38)}{\leqslant} \omega(f; (x_{i-1}, x_{i+1})) \underset{\substack{(23.36)\\(23.37)}}{\leqslant} \omega(f; [x_{i-1}, x_i]) + \omega(f; [x_i, x_{i+1}]) \equiv$$

$$\equiv \omega_i(f) + \omega_{i+1}(f) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

что противоречит условию (23.56).

Итак, в любом случае для каждой точки x, удовлетворяющей условиям (23.56), существует такой отрезок $[x_{i-1}, x_i]$ разбиения τ, что $x \in [x_{i-1}, x_i]$ и

$$\omega_i(f) \geqslant \frac{\varepsilon}{2}. \qquad (23.58)$$

Отберем все такие отрезки. Они покрывают множество X_ε, поэтому, согласно предварительному замечанию (в котором надо положить $\eta = \delta/2$), сумма их длин не меньше $\delta/2$. Следовательно, обозначив звездочкой у знака сигмы сумму, распространяемую только на отобранные отрезки, будем иметь

$$\Sigma^* \Delta x_i \geqslant \frac{\delta}{2}. \qquad (23.59)$$

Из неравенств (23.58) и (23.59) вытекает неравенство

$$\sum_{i=1}^{i_\tau} \omega_i(f)\Delta x_i \geqslant \Sigma^* \omega_i(f)\Delta x_i \underset{(23.58)}{\geqslant} \frac{\varepsilon}{2} \Sigma^* \Delta x_i \underset{(23.59)}{\geqslant} \frac{\delta\varepsilon}{4} > 0.$$

Так как числа δ и ε фиксированы, а τ — произвольное разбиение, то полученное неравенство противоречит, например, критерию интегрируемости (23.14) (см. п. 23.5), т. е. функция f не интегрируема на $[a, b]$.

Доказательство достаточности. Пусть для любых чисел $\varepsilon > 0$ и $\delta > 0$ множество $X_\varepsilon = [a, b]_\varepsilon$ можно покрыть конечной системой интервалов, сумма длин которых меньше δ. Функция f ограничена на отрезке $[a, b]$, поэтому существует такая постоянная $c > 0$, что

$$|f(x)| \leqslant c,\ x \in [a, b]. \tag{23.60}$$

Зададим произвольно $\varepsilon > 0$ и возьмем $\delta = \frac{\varepsilon}{4c}$. Существует конечная система интервалов (α_i, β_i), $i = 1, 2, \ldots, n$, покрывающих множество $X_{\frac{\varepsilon}{2(b-a)}}$, с суммой длин меньше $\frac{\varepsilon}{4c}$:

$$X_{\frac{\varepsilon}{2(b-a)}} \subset \bigcup_{i=1}^{n} (\alpha_i, \beta_i), \tag{23.61}$$

$$\sum_{i=1}^{n} (\beta_i - \alpha_i) < \frac{\varepsilon}{4c}. \tag{23.62}$$

Объединение всех соответствующих отрезков $[\alpha_i, \beta_i]$, $i = 1, 2, \ldots, n$, можно представить в виде объединения конечного множества отрезков $[\lambda_l, \mu_l]$, $l = 1, 2, \ldots, m$, с не пересекающимися попарно внутренностями и с концами λ_l, μ_l, равными либо α_i, либо β_i, $i = 1, 2, \ldots, n$, либо a, либо b. Тогда

$$\sum_{l=1}^{m} (\mu_l - \lambda_l) = \sum_{i=1}^{n} (\beta_i - \alpha_i) \underset{(23.62)}{<} \frac{\varepsilon}{4c}. \tag{23.63}$$

Так как

$$\omega(f; [\lambda_l, \mu_l]) \underset{(23.60)}{\leqslant} < 2c, \tag{23.64}$$

то

$$\sum_{l=1}^{m} \omega(f; [\lambda_l, \mu_l])(\mu_l - \lambda_l) \underset{(23.64)}{<} 2c \sum_{l=1}^{m} (\mu_l - \lambda_l) \underset{(23.63)}{<} \frac{\varepsilon}{2}. \tag{23.65}$$

Удалим из отрезка $[a, b]$ все точки, принадлежащие отрезкам $[\alpha_i, \beta_i]$. Оставшееся множество представляет собой объединение конечного множества промежутков с концами ξ_j и η_j, $\xi_j < \eta_j$, $j = 1, 2, \ldots, p$, среди которых может быть не

более двух полуинтервалов с концами $\xi_j = a$ или $\eta_j = b$, а все остальные являются интервалами. При этом, согласно включению (23.61), пересечение каждого из отрезков $[\xi_j, \eta_j]$ с множеством $X_{\frac{\varepsilon}{2(b-a)}}$ пусто, следовательно, в любой точке $x \in [\xi_j, \eta_j]$, $j = 1, 2, \ldots, p$, выполняется неравенство $\omega(f; x) < \frac{\varepsilon}{2(b-a)}$. Ясно также, что $\sum_{j=1}^{p}(\eta_j - \xi_j) \leqslant b - a$. Отсюда, в силу следствия леммы 5, вытекает, что для каждого отрезка $[\xi_j, \eta_j]$ существует такое его разбиение

$$\tau_j = \{\zeta_{k_j}\}_{k_j=0}^{k_j = k_{\tau_j}},$$

что

$$\sum_{k_j=1}^{i_{\tau_j}} \omega(f; [\zeta_{k_j-1}, \zeta_{k_j}](\zeta_{k_j} - \zeta_{k_j-1}) < \frac{\varepsilon}{2(b-a)}(\eta_j - \zeta_j). \qquad (23.66)$$

Пусть теперь $\tau = \{x_\nu\}_{\nu=0}^{\nu=\nu_\tau}$ — разбиение всего отрезка $[a, b]$, состоящее из всех точек λ_l, μ_l, $l = 1, 2, \ldots, m$, и точек ζ_{k_j}, $k_j = 1, 2, \ldots, k_{\tau_j}$, $j = 1, 2, \ldots, p$. Тогда

$$\sum_{\nu=1}^{\nu_\tau} \omega_\nu(f)\Delta x_\nu = \sum_{l=1}^{m} \omega(f; [\lambda_l, \mu_l])(\mu_l - \lambda_l) +$$
$$+ \sum_{j=1}^{p} \sum_{k_j=1}^{i_{\tau_j}} \omega(f; [\zeta_{k_j-1}; \zeta_{k_j}])(\zeta_{k_j} - \zeta_{k_j-1}) \underset{\substack{(23.65)\\(23.66)}}{<}$$
$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2(b-a)} \sum_{j=1}^{p}(\eta_j - \zeta_j) \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Согласно следствию из критерия интегрируемости Римана (см. теорему 8), отсюда вытекает, что функция f интегрируема на отрезке $[a, b]$. □

23.10*. Критерий интегрируемости Лебега

Докажем еще критерий интегрируемости функций, принадлежащий Лебегу[1]. Он прост по форме и удобен для приложений. При его формулировке придется иметь дело с бесконечными суммами положительных слагаемых. Определим это понятие.

[1] А. Лебег (1875—1941) — французский математик.

Если $a_n > 0$, $n = 1, 2, \ldots$, *то суммой* $\sum_{n=1}^{\infty} a_n$ *называется конечный или бесконечный предел* $\lim_{n \to \infty} (a_1 + \ldots + a_n)$.

Этот предел, конечный или равный $+\infty$, всегда существует, так как последовательность $\{a_1 + \ldots + a_n\}$, в силу условия $a_n > 0$, возрастает. Таким образом,

$$\sum_{n=1}^{\infty} a_n \overset{\text{def}}{=} \lim_{n \to \infty} \sum_{k=1}^{n} a_k.$$

Суммы бесконечного числа слагаемых будут подробно рассмотрены в главе III.

Определение 5. *Множество* X, *лежащее на числовой оси, называется множеством лебеговой меры нуль, если для любого* $\varepsilon > 0$ *существует покрытие этого множества конечной или счетной системой интервалов, сумма длин которых меньше* ε.

Пример. Всякое конечное или счетное множество является множеством лебеговой меры нуль.

Пусть $X = \{x_n\}$ — конечное или счетное множество (индекс n может быть либо любым натуральным числом, когда X — счетное множество, либо принимать только значения, не превосходящие некоторого фиксированного натурального числа, тогда множество X конечно). Зададим произвольно $\varepsilon > 0$. Система интервалов $\left(x_n - \frac{\varepsilon}{2^{n+2}}, x_n + \frac{\varepsilon}{2^{n+2}}\right)$, $n = 1, 2, \ldots$, очевидно, покрывает множество X, а сумма их длин меньше ε:

$$\sum_{n=1}^{\infty} a_n \frac{\varepsilon}{2^{n+1}} = \frac{\varepsilon}{4} \frac{1}{1 - 1/2} = \frac{\varepsilon}{2} < \varepsilon.$$

Например, множество всех рациональных чисел является множеством лебеговой меры нуль.

Задача 16. Построить пример несчетного множества лебеговой меры нуль.

ТЕОРЕМА 10 (теорема Лебега). *Для того чтобы ограниченная на отрезке функция была на нем интегрируема, необходимо и достаточно, чтобы множество ее точек разрыва было множеством лебеговой меры нуль.*

Доказательство необходимости. Пусть функция f интегрируема на отрезке $X = [a, b]$ и X_0 — множество ее точек разрыва. Зададим произвольно $\varepsilon > 0$. Согласно критерию интегрируемости Дюбуа-Реймона, для каждого $n = 1, 2, \ldots$

существует такая конечная система интервалов, покрывающая множество $X_{1/n}$ (см. (23.40)), сумма длин которых меньше $\varepsilon/2^n$. Объединение всех таких систем состоит из не более чем счетного множества интервалов, покрывающих, в силу формулы $X_0 = \bigcup_{n=1}^{\infty} X_{1/n}$ (см. (23.43)), все множество X_0 точек разрыва функции f. При этом сумма длин этих интервалов меньше, чем $\sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon$, т. е. X_0 действительно является множеством лебеговой меры нуль.

Доказательство достаточности. Пусть множество X_0 точек разрыва функции f, ограниченной на отрезке $X = [a, b]$, является множеством лебеговой меры нуль. Зададим произвольно $\varepsilon > 0$ и $\delta > 0$. Тогда существует не более чем счетная система интервалов, покрывающая множество X_0, с суммой длин интервалов меньшей δ. Выберем натуральное число n так, чтобы $1/n < \varepsilon$. Указанная выше система интервалов, являясь покрытием множества X_0, покрывает, в силу формулы $X_0 = \bigcup_{n=1}^{\infty} X_{1/n}$ (см. (23.43)), множество $X_{1/n}$, а следовательно, и множество X_ε, ибо $X_\varepsilon \subset X_{1/n}$ (см. (23.41)). Множество X_ε является ограниченным замкнутым множеством (см. следствие 2 леммы 3), поэтому из рассматриваемой системы интервалов, покрывающих его, можно выделить конечную систему интервалов, по-прежнему покрывающих множество X_ε (см. лемму 4 в п. 23.8), причем сумма длин входящих в нее интервалов (она не превосходит суммы длин всех интервалов исходной системы, покрывающей множество X_0) меньше δ. В силу критерия интегрируемости Дюбуа-Реймона, отсюда следует интегрируемость функции f. □

Из критерия интегрируемости Лебега следует, например, что всякая ограниченная на отрезке функция, имеющая конечное или счетное множество точек разрыва, является интегрируемой. Множество всех рациональных чисел счетно и, следовательно, является множеством лебеговой меры нуль, поэтому из критерия интегрируемости Лебега сразу следует интегрируемость функции Римана, рассмотренной в п. 23.6.

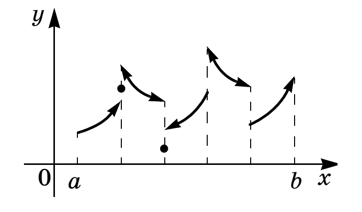


Рис. 110

В дальнейшем будем использовать интегрируемость так называемых кусочно-непрерывных функций. Определим эти функции.

Определение 6. *Функция называется кусочно-непрерывной на отрезке, если она задана и непрерывна во всех точках этого отрезка, кроме, быть может, конечного их множества, в точках которого имеет конечные односторонние пределы по проколотым окрестностям.*

В частности, кусочно-непрерывная на отрезке функция может быть не определена на концах этого отрезка, но она в левом конце отрезка имеет в указанном смысле конечный предел справа, а в правом конце — конечный предел слева (рис. 110).

Это означает, что если функция f кусочно-непрерывная на отрезке $[a, b]$, то существует такое разбиение $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ этого отрезка, что функция f непрерывна на каждом интервале (x_{i-1}, x_i) и существуют конечные пределы

$$f(x_{i-1}+0) = \lim_{x \to x_{i-1}+0} f(x)$$

и

$$f(x_i-0) = \lim_{x \to x_i-0} f(x),\ \ i = 1, 2, \dots, i_\tau.$$

Такой функции f бывает удобно сопоставить набор непрерывных функций f_i, заданных соответственно на отрезках $[x_{i-1}, x_i]$ равенствами

$$f_i(x) = \begin{cases} f(x) & \text{при } x_{i-1} < x < x_i, \\ f(x_{i-1}+0) & \text{при } x = x_{i-1}, \\ f(x_i-0) & \text{при } x = x_i. \end{cases} \tag{23.67}$$

Функция f_i действительно непрерывна на отрезке $[x_{i-1}, x_i]$, так как во всех внутренних точках этого отрезка она совпадает с непрерывной функцией f, а на его концах значения функции получены «непрерывным продолжением»: $f_i(x_{i-1}) = f(x_{i-1}+0)$ и $f_i(x_i) = f(x_i-0)$. Все функции f_i, будучи непрерывными на соответствующих отрезках, ограничены и этих

функций конечное множество, поэтому ограничена и функция f (множество ее значений может отличаться от объединения множества значений функций f_i, $i = 1, 2, \ldots, i_\tau$, только лишь на конечное множество значений функций f и f_i в точках $x_0, x_1, \ldots, x_{i_\tau}$).

Итак, всякая кусочно-непрерывная на отрезке функция ограничена. Кроме того она, очевидно, интегрируема на этом отрезке, так как ограничена и имеет конечное множество точек разрыва, а всякое конечное множество является, как это отмечено выше, множеством лебеговой меры нуль.

§ 24.

Свойства интегрируемых функций

24.1. Свойства определенного интеграла

В дальнейшем обозначения и терминологию, введенные в предыдущем параграфе, будем использовать, не делая специальных ссылок.

Прежде всего заметим, что интеграл от функции является числом, сопоставляемым заданной функции согласно данному выше определению, поэтому это число не зависит от выбора обозначения для аргумента подынтегральной функции, т. е. от обозначения *переменной интегрирования*:

$$\int_a^b f(x)\,dx = \int_a^b f(t)\,dt = \int_a^b f(\xi)\,d\xi. \qquad (24.1)$$

Перейдем теперь к рассмотрению **основных свойств определенного интеграла:**

1^0. $\int_a^b dx = b - a$.

Действительно, здесь подынтегральная функция равна единице, поэтому при любом разбиении $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ все интегральные суммы Римана σ_τ равны $b - a$:

$$\sigma_\tau = \sum_{i=1}^{i_\tau} \Delta x_i = b - a. \ \square$$

2^0. *Если функция f интегрируема на отрезке* $[a, b]$, *то она интегрируема на любом отрезке* $[a^*, b^*]$, *содержащемся в* $[a, b]$.

Доказательство. Прежде всего, если функция f ограничена на отрезке $[a, b]$, то она, очевидно, ограничена и на $[a^*, b^*]$. Далее, каково бы ни было разбиение $\tau^* = \{x_i^*\}_{i=0}^{i=i_{\tau^*}}$ отрезка $[a^*, b^*]$ мелкости δ_{τ^*}, его всегда можно продолжить в разбиение $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ отрезка $[a, b]$ такой же мелкости $|\tau| = |\tau^*|$; для этого достаточно добавить к точкам x_i^*, $i = 1, 2, \ldots$ $\ldots, i_{\tau^*}$ конечное число соответствующим образом выбранных точек, принадлежащих отрезку $[a, b]$, но не принадлежащих отрезку $[a^*, b^*]$.

Полагая

$$m_i^* = \inf_{x_{i-1}^* \leqslant x \leqslant x_i^*} f(x), \quad M_i^* = \sup_{x_{i-1}^* \leqslant x \leqslant x_i^*} f(x),$$
$$\Delta x_i^* = x_i^* - x_{i-1}^*, \quad i = 1, 2, \ldots, i_{\tau^*}, \tag{24.2}$$

и, как обычно,

$$m_i = \inf_{x_{i-1} \leqslant x \leqslant x_i} f(x), \quad M_i = \sup_{x_{i-1} \leqslant x \leqslant x_i} f(x),$$
$$\Delta x_i = x_i - x_{i-1}, \quad i = 1, 2, \ldots, i_\tau, \tag{24.3}$$

замечая, что каждое слагаемое суммы $\sum_{i=1}^{i_{\tau^*}} (M_i^* - m_i^*)\Delta x_i^*$ является и слагаемым суммы $\sum_{i=1}^{i_{\tau^*}} (M_i - m_i)\Delta x_i$ и что все слагаемые обеих сумм неотрицательны, имеем

$$0 \leqslant S_{\tau^*} - s_{\tau^*} = \sum_{i=1}^{i_{\tau^*}} (M_i^* - m_i^*)\Delta x_i^* \leqslant \sum_{i=1}^{i_\tau} (M_i - m_i)\Delta x_i = S_\tau - s_\tau. \tag{24.4}$$

Если функция f интегрируема на отрезке $[a, b]$, то (см. п. 23.5)

$$\lim_{|\tau| \to 0} (S_\tau - s_\tau) = 0. \tag{24.5}$$

Поскольку $|\tau| = |\tau^*|$, из (24.5) и из неравенства (24.4) имеем

$$\lim_{|\tau^*| \to 0} (S_{\tau^*} - s_{\tau^*}) = 0, \tag{24.6}$$

т. е. (см. п. 23.5) функция f интегрируема на отрезке $[a^*, b^*]$. □

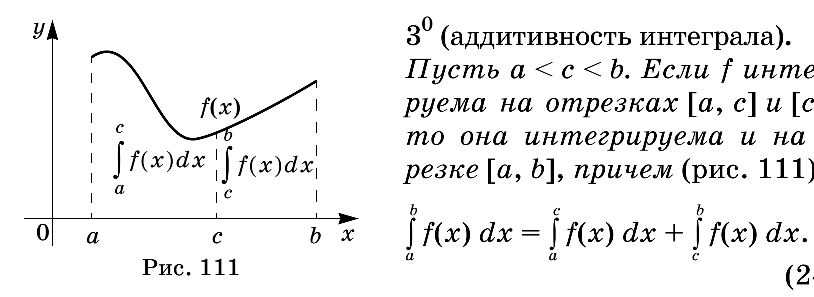


Рис. 111

3^0 (аддитивность интеграла). *Пусть $a < c < b$. Если f интегрируема на отрезках $[a, c]$ и $[c, b]$, то она интегрируема и на отрезке $[a, b]$, причем* (рис. 111)

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx. \qquad (24.7)$$

Доказательство. Интегрируемость функции f на отрезке $[a, b]$ при условии, что она интегрируема на отрезках $[a, c]$ и $[c, b]$, доказана в п. 23.7* (см. замечание 1). Докажем формулу (24.7). Пусть τ_1 и τ_2 — какие-либо разбиения соответственно отрезков $[a, c]$ и $[c, b]$, а τ — разбиение отрезка $[a, b]$, получающееся объединением разбиений τ_1 и τ_2. Тогда, очевидно,

$$|\tau_1| \leqslant |\tau|, \; |\tau_2| \leqslant |\tau|. \qquad (24.8)$$

Если σ_{τ_1} и σ_{τ_2} — какие-либо суммы Римана функции f, соответствующие разбиениям τ_1 и τ_2, то

$$\sigma_\tau = \sigma_{\tau_1} + \sigma_{\tau_2} \qquad (24.9)$$

является некоторой суммой Римана функции f, соответствующей разбиению τ. В силу интегрируемости функции f на отрезках $[a, b]$, $[a, c]$ и $[c, b]$ существуют конечные пределы

$$\lim_{|\tau| \to 0} \sigma_\tau = \int_a^b f(x)\,dx, \quad \lim_{|\tau_1| \to 0} \sigma_{\tau_1} = \int_a^c f(x)\,dx, \quad \lim_{|\tau_2| \to 0} \sigma_{\tau_2} = \int_c^b f(x)\,dx.$$

Поэтому, перейдя к пределу при $|\tau| \to 0$ в равенстве (24.9), получим, в силу (24.8), формулу (24.7). □

Замечание. Согласно определению интеграла $\int_a^b f(x)\,dx$ при $b \leqslant a$ (см. п. 23.1) формула (24.7) остается в силе и при $c > b$, если только функция f интегрируема на отрезке $[a, c]$.

В самом деле, если $c > b$, то по доказанному (для простоты опускаем записи подынтегральных выражений):

$$\int_a^b + \int_b^c = \int_a^c \text{ и, следовательно, } \int_a^b = \int_a^c - \int_b^c = \int_a^c + \int_c^b.$$

4^0. *Если функции f и g интегрируемы на отрезке $[a, b]$, то их сумма $f + g$ также интегрируема на нем, причем*

$$\int_a^b [f(x) + g(x)]\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx. \qquad (24.10)$$

Доказательство. В самом деле, каковы бы ни были разбиение $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ отрезка $[a, b]$ и точки $\xi_i \in [x_{i-1}, x_i]$, $i = 1, 2, \ldots, i_\tau$, имеем

$$\sigma_\tau(f + g) = \sum_{i=1}^{i_\tau} [f(\xi_i) + g(\xi_i)]\Delta x_i =$$
$$= \sum_{i=1}^{i_\tau} f(\xi_i)\Delta x_i + \sum_{i=1}^{i_\tau} g(\xi_i)\Delta x_i = \sigma_\tau(f) + \sigma_\tau(g). \qquad (24.11)$$

В силу интегрируемости функций f и g, существуют пределы интегральных сумм $\sigma_\tau(f)$ и $\sigma_\tau(g)$ при $|\tau| \to 0$, поэтому из (24.11) следует, что существует и предел (почему?) интегральной суммы $\sigma_\tau(f + g)$, причем

$$\lim_{|\tau| \to 0} \sigma_\tau(f + g) = \lim_{|\tau| \to 0} \sigma_\tau(f) + \lim_{|\tau| \to 0} \sigma_\tau(g), \qquad (24.12)$$

что и означает интегрируемость функции $f + g$ на отрезке $[a, b]$.

Согласно же определению интеграла,

$$\lim_{|\tau| \to 0} \sigma_\tau(f + g) = \int_a^b [f(x) + g(x)]\,dx,$$

$$\lim_{|\tau| \to 0} \sigma_\tau(f) = \int_a^b f(x)\,dx, \quad \lim_{|\tau| \to 0} \sigma_\tau(g) = \int_a^b g(x)\,dx.$$

Подставляя эти выражения в формулу (24.12), получим (24.10). □

5^0. *Пусть функция f интегрируема на отрезке $[a, b]$ и c — постоянная; тогда функция cf также интегрируема на этом отрезке и*

$$\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx. \qquad (24.13)$$

Доказательство. Каковы бы ни были разбиение $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ отрезка $[a, b]$ и точки $\xi_i \in [x_{i-1}, x_i]$, $i = 1, 2, \ldots, i_\tau$, имеем

$$\sigma_\tau(cf) = \sum_{i=1}^{i_\tau} cf(\xi_i)\Delta x_i = c\sum_{i=1}^{i_\tau} f(\xi_i)\Delta x_i = c\sigma_\tau(f),$$

отсюда

$$\int_a^b cf(x)\,dx = \lim_{|\tau| \to 0} \sigma_\tau(cf) = \lim_{|\tau| \to 0} c\sigma_\tau(f) =$$

$$= c \lim_{|\tau| \to 0} \sigma_\tau(f) = c\int_a^b f(x)\,dx. \ \square$$

Из последних двух свойств вытекает следствие: *если каждая из функций* f_i, $i = 1, 2, \ldots, n$, *интегрируема на отрезке* $[a, b]$, *а* λ_i *— произвольные постоянные, то функция* $\sum_{i=1}^{n} \lambda_i f_i$ *интегрируема на* $[a, b]$, *причем*

$$\int_a^b \sum_{i=1}^{n} (\lambda_i f_i(x)\,dx = \sum_{i=1}^{n} \lambda_i \int_a^b f_i(x)\,dx. \qquad (24.14)$$

Это свойство определенного интеграла называется его *линейностью*.

6^0. *Пусть функции* $f(x)$ *и* $g(x)$ *интегрируемы на отрезке* $[a, b]$. *Тогда и их произведение* $f(x)g(x)$ *интегрируемо на нем.*

Доказательство. В силу интегрируемости функций f и g на отрезке $[a, b]$, они ограничены на этом отрезке, т. е. существуют постоянные $A > 0$ и $B > 0$ такие, что

$$|f(x)| \leqslant A, \ |g(x)| \leqslant B \qquad (24.15)$$

для всех $x \in [a, b]$. Поэтому произведение $f(x)g(x)$ также ограничено: для всех точек $x \in [a, b]$ выполняется неравенство

$$|f(x)g(x)| \leqslant AB. \qquad (24.16)$$

Пусть $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ — какое-либо разбиение отрезка $[a, b]$. Оценим выражение $f(x'')g(x'') - f(x')g(x')$, x', $x'' \in [a, b]$; для этого добавим и вычтем из него $f(x')g(x'')$:

$$f(x'')g(x'') - f(x')g(x') =$$

$$= [f(x'') - f\{x')]g(x'') + [g(x'') - g(x')]f(x'). \qquad (24.17)$$

Для точек $x' \in [x_{i-1}, x_i]$ и $x'' \in [x_{i-1}, x_i]$ из (24.15) и (24.17) следует, что

$$|f(x'')g(x'') - f(x')g(x')| \leqslant B\omega_i(f) + A\omega_i(g), \qquad (24.18)$$

где $\omega_i(f)$ и $\omega_i(g)$ — колебания функций f и g на отрезках $[x_{i-1}, x_i]$, $i = 1, 2, \ldots, i_\tau$.

Из неравенства (24.18) для колебания $\omega_i(fg)$ произведения fg на отрезке $[x_{i-1}, x_i]$ вытекает оценка

$$\omega_i(fg) \leqslant B\omega_i(f) + A\omega_i(g). \qquad (24.19)$$

Отсюда

$$\sum_{i=1}^{i_\tau} \omega_i(fg)\Delta x_i \leqslant B\sum_{i=1}^{i_\tau} \omega_i(f)\Delta x_i + A\sum_{i=1}^{i_\tau} \omega_i(g)\Delta x_i. \qquad (24.20)$$

В силу интегрируемости функций f и g (см. следствие 2 из теоремы 3 в п. 23.5),

$$\lim_{|\tau| \to 0} \sum_{i=1}^{i_\tau} \omega_i(f)\Delta x_i = \lim_{|\tau| \to 0} \sum_{i=1}^{i_\tau} \omega_i(g)\Delta x_i = 0.$$

Поэтому из оценки (24.20) следует равенство

$$\lim_{|\tau| \to 0} \sum_{i=1}^{i_\tau} \omega_i(fg)\Delta x_i = 0,$$

которое в силу того же следствия из теоремы 3 и влечет за собой интегрируемость произведения fg на отрезке $[a, b]$. □

Методом математической индукции легко доказать, что если каждая из функций $f_i(x)$, $i = 1, 2, \ldots, n$, интегрируема на отрезке $[a, b]$, то и их произведение интегрируемо на этом отрезке. В частности, вместе с функцией $f(x)$ интегрируема и ее степень $[f(x)]^n$ при любом натуральном n.

7^0. *Если функция $f(x)$ интегрируема на отрезке $[a, b]$ и нижняя грань функции $|f(x)|$ на нем положительна, то и $\frac{1}{f(x)}$ интегрируема на этом отрезке.*

Доказательство. Если $\inf\limits_{[a,b]} |f(x)| = m > 0$, то $|f(x)| \geqslant m$ всюду на $[a, b]$, следовательно, $\frac{1}{|f(x)|} \leqslant \frac{1}{m}$ для всех $x \in [a, b]$; поэтому

$$\left|\frac{1}{f(x_2)} - \frac{1}{f(x_1)}\right| \leqslant \frac{|f(x_1) - f(x_2)|}{m^2}$$

при любых $x_1, x_2 \in [a, b]$. Отсюда следует, что если $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ — произвольное разбиение отрезка $[a, b]$, то

$$\omega_i\left(\frac{1}{f}\right) \leqslant \frac{1}{m^2}\omega_i(f),$$

следовательно,

$$0 \leqslant \lim_{|\tau| \to 0} \sum_{i=1}^{i_\tau} \omega_i\left(\frac{1}{f}\right)\Delta x_i \leqslant \frac{1}{m^2} \lim_{|\tau| \to 0} \sum_{i=1}^{i_\tau} \omega_i(f)\Delta x_i = 0,$$

поэтому $\lim\limits_{|\tau| \to 0} \sum\limits_{i=1}^{i_\tau} \omega_i\left(\frac{1}{f}\right)\Delta x_i = 0$. В силу следствия 2 теоремы 3 функция $\frac{1}{f}$ интегрируема на отрезке $[a, b]$. □

СЛЕДСТВИЕ. *Если функции f и g интегрируемы на отрезке $[a, b]$ и нижняя грань функции $|g|$ положительна, то и частное $\frac{f}{g}$ интегрируемо на этом отрезке.*

Утверждение вытекает, в силу свойств 6^0 и 7^0, из того, что $\frac{f}{g} = f \cdot \frac{1}{g}$. □

8^0. *Если функция f неотрицательна и интегрируема на отрезке $[a, b]$, то*

$$\int_a^b f(x)\,dx \geqslant 0. \tag{24.21}$$

Доказательство. В самом деле, каковы бы ни были разбиение $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ отрезка $[a, b]$ и точки $\xi_i \in [x_{i-1}, x_i]$, $i = 1, 2, \ldots, i_\tau$, для функции $f \geqslant 0$ имеем

$$\sigma_\tau(f) = \sum_{i=1}^{i_\tau} f(\xi_i)\Delta x_i \geqslant 0. \tag{24.22}$$

Если функция f интегрируема на отрезке $[a, b]$, то, переходя к пределу в (24.22) при $|\tau| \to 0$, получим неравенство (24.21). □

СЛЕДСТВИЕ. *Если функции f и g интегрируемы на отрезке $[a, b]$ и для всех $x \in [a, b]$ выполняется неравенство*

$$f(x) \geqslant g(x), \tag{24.23}$$

то

$$\int_a^b f(x)\,dx \geqslant \int_a^b g(x)\,dx. \tag{24.24}$$

Если интегрируемые функции f и g удовлетворяют неравенству (24.23), то

$$f(x) - g(x) \geqslant 0,\ x \in [a, b];$$

поэтому, замечая, что на основании следствия из свойств 4^0 и 5^0 функция $f - g$ интегрируема, в силу неравенства (24.21) имеем

$$\int_a^b [f(x) - g(x)]\, dx \geqslant 0.$$

Но (см. указанное выше следствие)

$$\int_a^b [f(x) - g(x)]\, dx = \int_a^b f(x)dx - \int_a^b g(x)\, dx$$

и, значит,

$$\int_a^b f(x)\, dx - \int_a^b g(x)\, dx \geqslant 0. \square$$

Доказанное утверждение означает, что обе части неравенства вида (24.23) можно интегрировать по одному и тому же промежутку. (В связи с этим заметим, что при дифференцировании обеих частей неравенства оно, вообще говоря, не сохраняется.)

9^0. *Пусть функция f интегрируема на отрезке $[a, b]$. Если она неотрицательна на нем: $f(x) \geqslant 0$, $x \in [a, b]$ — и существует точка $x_0 \in [a, b]$, в которой функция f непрерывна и положительна: $f(x_0) > 0$, то*

$$\int_a^b f(x)\, dx > 0. \qquad (24.25)$$

СЛЕДСТВИЕ. *Если функции f и g интегрируемы на отрезке $[a, b]$, для всех точек $x \in [a, b]$ выполняется неравенство*

$$f(x) \geqslant g(x),$$

существует такая точка $x_0 \in [a, b]$, что в ней обе функции f, g непрерывны, и имеет место неравенство

$$f(x_0) > g(x_0),$$

то справедливо неравенство

$$\int_a^b f(x)\, dx > \int_a^b g(x)\, dx. \qquad (24.26)$$

Доказательство свойства 9^0. Из очевидного неравенства $f(x_0) > \frac{f(x_0)}{2} > 0$ следует, что существует такое $\delta > 0$, что для всех $x \in [a, b] \cap U(x_0, \delta)$ выполняется неравенство

$f(x) \geqslant \frac{f(x_0)}{2}$ (см. свойство 2^0 пределов функций в п. 5.10). Пусть $[\alpha, \beta] \subset U(x_0, \delta) \cap [a, b]$, $\alpha < \beta$; тогда

$$\int_a^b f(x)\,dx \geqslant \int_\alpha^\beta f(x)\,dx \geqslant \frac{f(x_0)}{2}(\beta - \alpha) > 0. \ \square$$

Отметим, что если отказаться от условия непрерывности функции f в точке x_0, то может случиться, что для интегрируемой неотрицательной на отрезке функции, положительной в некоторой точке, интеграл по всему отрезку равен нулю. Так, например, функция

$$f(x) = \begin{cases} 0 \text{ при } 0 < x \leqslant 1, \\ 1 \text{ при } x = 0 \end{cases}$$

интегрируема и неотрицательна, $f(0) > 0$, но $\int_0^1 f(x)\,dx = 0$. Это равенство легко следует из определения интеграла.

Следствие свойства 9^0 доказывается аналогично следствию свойства 8^0.

10^0. *Если функция f интегрируема на отрезке $[a, b]$, то функция $|f|$ также интегрируема на нем и*

$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b |f(x)|\,dx, \ a < b. \qquad (24.27)$$

Действительно, во-первых, из ограниченности функции f, очевидно, следует и ограниченность функции $|f|$, а во-вторых, для любых двух точек $\xi \in [a, b]$ и $\eta \in [a, b]$ имеет место неравенство

$$||f(\xi)| - |f(\eta)|| \leqslant |f(\xi) - f(\eta)|,$$

откуда следует, что, каково бы ни было разбиение $\tau = \{x_i\}_{i=0}^{i=i\tau}$ отрезка $[a, b]$, обозначая через $\omega_i(f)$ и $\omega_i(|f|)$ соответственно колебания функций f и $|f|$ на отрезке $[x_{i-1}, x_i]$, получим

$$\omega_i(|f|) = \sup_{x, x' \in [x_{i-1}, x_i]} ||f(x')| - |f(x)|| \leqslant$$
$$\leqslant \sup_{x, x' \in [x_{i-1}, x_i]} |f(x') - f(x)| \leqslant \omega_i(f),$$

поэтому

$$0 \leqslant \sum_{i=1}^{i_\tau} \omega_i(|f|)\Delta x_i \leqslant \sum_{i=1}^{i_\tau} \omega_i(f)\Delta x_i.$$

Отсюда следует, что если

$$\lim_{|\tau| \to 0} \sum_{i=1}^{i_\tau} \omega_i(f)\Delta x_i = 0, \text{ то и } \lim_{|\tau| \to 0} \sum_{i=1}^{i_\tau} \omega_i(|f|)\Delta x_i = 0.$$

Это означает (см. п. 23.5), что из интегрируемости функции f следует интегрируемость функции $|f|$.

Пусть теперь $\xi_i \in [x_{i-1}, x_i]$, $i = 1, 2, \ldots, i_\tau$, тогда

$$|\sigma_\tau(f)| = \left|\sum_{i=1}^{i_\tau} f(\xi_i)\Delta x_i\right| \leqslant \sum_{i=1}^{i_\tau} |f(\xi_i)|\Delta x_i = \sigma_\tau(|f|).$$

Переходя в этом неравенстве к пределу при $|\tau| \to 0$ и замечая, что

$$\lim_{|\tau| \to 0} |\sigma_\tau(f)| = |\lim_{|\tau| \to 0} \sigma_\tau(f)| = \left|\int_a^b f(x)\,dx\right|,$$

$$\lim_{|\tau| \to 0} \sigma_\tau(|f|) = \int_a^b |f(x)|\,dx,$$

получим неравенство (24.27). □

Если отказаться от ограничения $a < b$, т. е. допускать случаи $a = b$ и $a > b$, то аналог неравенства (24.27) имеет вид

$$\left|\int_a^b f(x)\,dx\right| \leqslant \left|\int_a^b |f(x)|\,dx\right|. \qquad (24.28)$$

В самом деле, пусть $a < b$. Поскольку (см. свойство 8^0)

$$\left|\int_a^b |f(x)|\,dx\right| = \int_a^b |f(x)|\,dx,$$

неравенство (24.28) совпадает в этом случае с неравенством (24.27). Если же $a > b$, то, используя свойство (23.3) и неравенство (24.27), получим

$$\left|\int_a^b f(x)\,dx\right| = \left|\int_b^a f(x)\,dx\right| \leqslant \int_a^b |f(x)|\,dx = \left|\int_b^a |f(x)|\,dx\right| = \left|\int_a^b |f(x)|\,dx\right|.$$

Наконец, при $a = b$ неравенство (24.28) очевидно.

Примеры. 1. Пусть функция f — *четная* на отрезке $[-a, a]$ и интегрируемая на отрезке $[0, a]$. Тогда она интегрируема и на отрезке $[-a, a]$, причем

$$\int_{-a}^a f(x)\,dx = 2\int_0^a f(x)\,dx. \qquad (24.29)$$

Докажем, что функция f интегрируема на отрезке $[-a, 0]$ и что

$$\int_0^a f(x)\,dx, \qquad (24.30)$$

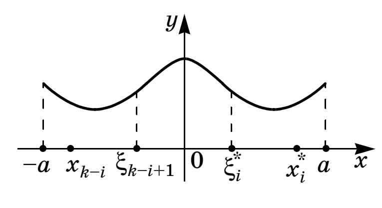


Рис. 112

отсюда, в силу аддитивности интеграла (свойство 3^0), и следует сразу формула (24.29), так как, опуская обозначение подынтегральной функции,

$$\int_{-a}^{a} = \int_{-a}^{0} + \int_{0}^{a} = 2\int_{0}^{a}.$$

Равенство (24.30) следует из того, что если f — четная функция, то преобразование симметрии числовой оси относительно нуля переводит ее интегральные суммы на отрезке $[0, a]$ в равные интегральные суммы по отрезку $[-a, 0]$ и наоборот. В самом деле, если $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ — разбиение отрезка $[-a, 0]$, то $\tau^* = \{x_i^*\}_{i=0}^{i=i_\tau}$, где $x_i^* = -x_{i_\tau - i}$, $i = 0, 1, 2, \ldots, i_\tau$ является разбиением отрезка $[0, a]$, причем мелкости обоих разбиений, очевидно, совпадают: $|\tau| = |\tau^*|$. Если для каждой точки $\xi_i \in [x_{i-1}, x_i]$ положить $\xi_i^* = -\xi_{i_\tau - i + 1}$, $i = 1, 2, \ldots, i_\tau$, то, в силу четности функции f, получим $f(\xi_i^*) = f(\xi_{i_\tau - i + 1})$ (рис. 112) и, следовательно, всякой интегральной сумме Римана $\sigma_\tau = \sum\limits_{i=1}^{i_\tau} f(\xi_i)\Delta x_i$ функции f на отрезке $[-a, 0]$ будет соответствовать равная ей интегральная сумма $\sigma_{\tau^*} = \sum\limits_{i=1}^{i_\tau} f(\xi_i^*)\Delta x_i^* = \sigma_\tau$ той же функции f, но на отрезке $[0, a]$ (здесь, как всегда, $\Delta x_i = x_i - x_{i-1}$, $\Delta x_i^* = x_i^* - x_{i-1}^*$ и легко видеть, что $\Delta x_i = \Delta x^*_{i_\tau - i + 1}$, $i = 1, 2, \ldots, i_\tau$). Поэтому

$$\lim_{|\tau| \to 0} \sigma_\tau = \lim_{|\tau^*| \to 0} \sigma_{\tau^*} = \int_0^a f(x)\,dx.$$

Таким образом, предел в левой части этого равенства существует, а это означает, что функция f интегрируема на отрезке $[-a, 0]$. Поскольку же указанный предел равен интегралу $\int_{-a}^{0} f(x)\,dx$, равенство (24.30) доказано. □

Рассмотренный пример можно обобщить. Если функция f интегрируема на отрезке $[a, b]$, $0 \leqslant a < b$, и

$$f^*(x) \overset{\text{def}}{=} f(-x), \quad -b \leqslant x \leqslant -a$$

(рис. 113), то функция f^* интегрируема на отрезке $[-b, -a]$ и

$$\int_{-a}^{-b} f^*(x)\,dx = \int_{a}^{b} f(x)\,dx. \qquad (24.31)$$

Это утверждение доказывается аналогично рассмотренному выше случаю.

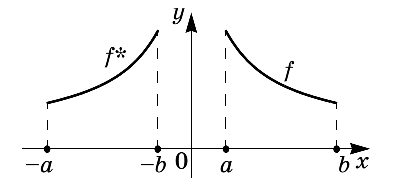


Рис. 113

УПРАЖНЕНИЕ 1. Доказать, что если нечетная функция f интегрируема на отрезке $[-a, a]$, то $\int_{-a}^{a} f(x)\,dx = 0$.

2. Рассмотрим теперь периодические функции.

Функция $f\colon X \to \boldsymbol{R}$, $X \subset \boldsymbol{R}$, называется *периодической на множестве X с периодом $T > 0$, если для любого $x \in X$ выполняется включение $x \pm T \in X$ и равенство $f(x + T) = f(x)$.*

Например, для функции $\sin x$ периодом является любое число, целочисленно кратное 2π, т. е. число вида $2\pi n$, $n = 0, \pm 1, \pm 2, \ldots$. Для постоянной функции периодом является любое положительное число.

УПРАЖНЕНИЯ. 2. Доказать, что у функции Дирихле (см. п. 5.2) любое положительное рациональное число является периодом, а любое положительное иррациональное число не является таковым.

3. Доказать, что функция $\sin x + \operatorname{tg} x$ имеет наименьший период, и найти его.

4. Привести пример двух функций, имеющих наименьший период, сумма которых не имеет наименьшего периода.

5. Доказать, что всякая непрерывная периодическая на всей числовой оси $\boldsymbol{R}$ функция ограничена на $\boldsymbol{R}$.

6. Доказать, что всякая непрерывная периодическая на всей числовой оси $\boldsymbol{R}$ функция равномерно непрерывна на $\boldsymbol{R}$.

Если функция f имеет период $T > 0$ и для некоторого числа a интегрируема на отрезке $[a, a + T]$, то, каково бы ни было число b, она интегрируема на отрезке $[b, b + T]$ и имеет место равенство

$$\int_{b}^{b+T} f(x)\,dx = \int_{a}^{a+T} f(x)\,dx, \qquad (24.32)$$

т. е. интеграл от периодической функции по отрезку, равному по длине периоду, не зависит от положения этого отрезка на числовой прямой (рис. 114).

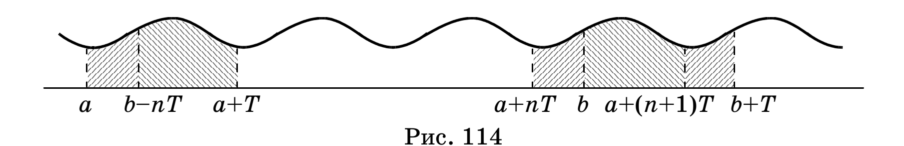


Рис. 114

Докажем это. Пусть для определенности $b \geqslant a$, тогда существует такое неотрицательное целое n, что

$$a + nT \leqslant b < a + (n + 1)T,$$

и, следовательно,

$$a \leqslant b - nT < a + T, \ b < a + (n + 1)T \leqslant b + T.$$

Заметим, что если функция f интегрируема на некотором отрезке $[c, d]$, то для любого неотрицательного целого n функция f интегрируема на отрезке $[c + nT, d + nT]$ и справедливо равенство

$$\int_c^d f(x)\,dx = \int_{c+nT}^{d+nT} f(x)\,dx. \tag{24.33}$$

Действительно, в силу периодичности функции f, при сдвиге отрезка $[c, d]$ в отрезок $[c + nT, d + nT]$, т. е. при преобразовании аргумента $x' = x + nT$, в соответствующих друг другу точках этих отрезков функция f принимает одинаковые значения. В силу этого, можно тем же методом, который был применен в предыдущем примере, используя только вместо симметрии сдвиг, легко показать, что функция f интегрируема на отрезке $[c + nT, d + nT]$ и что имеет место равенство (24.33).

Применив этот результат к отрезкам $[a, b - nT]$ и $[b - nT, a + T]$, прибавив в первом случае к обоим концам отрезка число $(n + 1)T$, а во втором — число nT, получим, во-первых, что функция f интегрируема на отрезках $[a + (n + 1)T, b + T]$ и $[b, a + (n + 1)T]$, и, во-вторых, что

$$\int_a^{b-nT} f(x)\,dx = \int_{a+(n+1)T}^{b+T} f(x)\,dx,$$

$$\int_{b-nT}^{a-T} f(x)\,dx = \int_b^{a+(n+1)T} f(x)\,dx. \tag{24.34}$$

Складывая эти равенства, в силу свойства аддитивности интеграла (см. свойство 3^0), получим формулу (24.32). Аналогично рассматривается случай $b < a$. □

Мы еще вернемся к формулам (24.31) и (24.32) в разделе о замене переменного в интеграле (см. п. 26.1).

24.2. Первая теорема о среднем значении для определенного интеграла

ТЕОРЕМА 1. *Пусть*:
1) *функции* f *и* g: *интегрируемы на отрезке* $[a, b]$;
2) $m \leqslant f(x) \leqslant M$, $x \in [a, b]$;
3) *функция* g *не меняет знака на отрезке* $[a, b]$, *т. е. либо неотрицательна, либо неположительна на нем. Тогда существует такое число* μ, *что*

$$m \leqslant \mu \leqslant M \tag{24.36}$$

и

$$\int_a^b f(x)g(x)\,dx = \mu\int_a^b g(x)\,dx. \tag{24.37}$$

СЛЕДСТВИЕ. *При дополнительном предположении непрерывности функции* f *на отрезке* $[a, b]$ *существует такая точка* ξ *на интервале* (a, b), *что*

$$\int_a^b f(x)g(x)\,dx = f(\xi)\int_a^b g(x)\,dx. \tag{24.38}$$

В частности, при $g(x) = 1$ *на отрезке* $[a, b]$

$$\int_a^b f(x)\,dx = f(\xi)(b - a). \tag{24.39}$$

Последняя формула в случае неотрицательной на отрезке $[a, b]$ функции f имеет простой геометрический смысл: площадь криволинейной трапеции, образованной графиком функции f, равна площади прямоугольника с основанием длины $b - a$ и высотой длины $f(\xi)$ (рис. 115).

Доказательство теоремы. Умножая неравенство $m \leqslant f(x) \leqslant M$ на $g(x)$, при $g(x) \geqslant 0$ получаем

$$mg(x) \leqslant f(x)g(x) \leqslant Mg(x),$$

а при $g(x) \leqslant 0$

$$mg(x) \geqslant f(x)g(x) \geqslant Mg(x).$$

Интегрируя эти неравенства на основании следствия из свойств 8^0 (п. 24.1), будем иметь

$$m\int_a^b g(x)\,dx \leqslant \int_a^b f(x)g(x)\,dx \leqslant$$

$$\leqslant M\int_a^b g(x)\,dx, \tag{24.40}$$

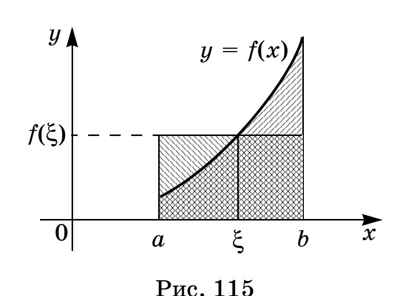


Рис. 115

соответственно

$$m\int_a^b g(x)\,dx \geqslant \int_a^b f(x)g(x)\,dx \geqslant M\int_a^b g(x)\,dx. \qquad (24.41)$$

Если $\int_a^b g(x)\,dx = 0$, то как в первом, так и во втором случаях

$$\int_a^b f(x)g(x)\,dx = 0.$$

Таким образом, если $\int_a^b g(x)\,dx = 0$, то обе части равенства (24.37) при любом μ обращаются в нуль, т. е. при выполнении условия $\int_a^b g(x)\,dx = 0$ равенство (24.37) справедливо при любом выборе числа μ, в частности и при $m \leqslant \mu \leqslant M$.

Если же $\int_a^b g(x)\,dx \neq 0$, то при $g(x) \geqslant 0$, $x \in [a, b]$, имеем $\int_a^b g(x)\,dx > 0$, а при $g(x) \leqslant 0$, $x \in [a, b]$, соответственно $\int_a^b g(x)\,dx < < 0$. Разделив неравенства (24.40) и (24.41) на интеграл $\int_a^b g(x)\,dx$, получим в обоих случаях одно и то же неравенство

$$m \leqslant \frac{\int_a^b f(x)g(x)dx}{\int_a^b g(x)dx} \leqslant M. \qquad (24.42)$$

Полагая

$$\mu \overset{\text{def}}{=} \frac{\int_a^b f(x)dx}{\int_a^b g(x)dx}, \qquad (24.43)$$

видим, что при таком выборе μ выполняются как условие (24.36) (в силу (24.42)), так и (24.37) (в силу (24.43)).

Доказательство следствия. Если функция f непрерывна на отрезке $[a, b]$, то, согласно теореме Вейерштрасса (п. 6.1), она достигает своего наибольшего и наименьшего значения в некоторых точках α и β этого отрезка:

$$f(\alpha) = \min_{[a,b]} f(x), \quad f(\beta) = \max_{[a,b]} f(x). \qquad (24.44)$$

При

$$m = f(\alpha), \quad M = f(\beta), \qquad (24.45)$$

выполняется условие (24.35) теоремы и, следовательно, существует такое число μ,

$$m \leqslant \mu \leqslant M, \qquad (24.46)$$

для которого выполняется равенство (24.37).

В силу условий (24.45) и (24.46), согласно теореме Больцано—Коши (см. п. 6.2) о промежуточных значениях непрерывной функции, на отрезке $[\alpha, \beta] \subset [a, b]$ существует точка ξ, для которой имеет место равенство $f(\xi) = \mu$ (рис. 116), поэтому и равенство (24.38). Покажем, что, более того, точку ξ можно выбрать всегда так, что она будет лежать не только на отрезке $[a, b]$, но и являться его внутренней точкой.

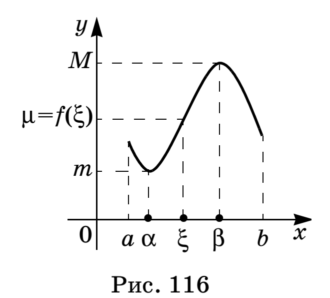


Рис. 116

Если $\int_a^b g(x)\,dx = 0$, то из формулы (24.37) следует, что

$$\int_a^b f(x)g(x)\,dx = 0.$$

Поэтому равенство (24.38) выполняется при любом выборе точки $\xi \in (a, b)$.

Пусть теперь

$$\int_a^b g(x)\,dx \neq 0 \tag{24.47}$$

и для определенности $g(x) \geqslant 0$ во всех точках x отрезка $[a, b]$, следовательно,

$$\int_a^b g(x)\,dx \geqslant 0. \tag{24.48}$$

(Случай $g(x) \leqslant 0$, $a \leqslant x \leqslant b$, сводится к рассматриваемому заменой функции $g(x)$ на функцию $-g(x)$: применив к неотрицательной функции $-g(x)$ формулу (24.38) и умножив обе части равенства на -1, получим и в этом случае формулу (24.38).)

Из выполнения условий (24.47) и (24.48) следует, что

$$\int_a^b g(x)\,dx > 0. \tag{24.49}$$

В силу неравенства (24.46) возможны три случая: $m < \mu < M$, $\mu = M$ и $\mu = m$. Если $m < \mu < M$, то в силу условий (24.45), согласно теореме Больцано—Коши о промежуточных значениях непрерывных функций между точками α и β, т. е. на интервале (a, b), существует такая точка ξ, что $f(\xi) = \mu$.

Если же $\mu = M$ (случай $\mu = m$ рассматривается аналогично), то равенство (24.38) принимает вид

$$\int_a^b f(x)g(x)\,dx = M\int_a^b g(x)\,dx,$$

откуда

$$\int_a^b [M - f(x)]g(x)\,dx = 0. \qquad (24.50)$$

Покажем, что существует такая точка $\xi \in (a, b)$, что $f(\xi) = M$. Предварительно заметим, что

$$\int_a^b g(x)\,dx = \lim_{\varepsilon \to +0} \int_{a+\varepsilon}^{b-\varepsilon} g(x)\,dx. \qquad (24.51)$$

В самом деле, функция $g(x)$ интегрируема на отрезке $[a, b]$, поэтому и ограничена на нем, т. е. существует такая постоянная $c > 0$, что для всех точек $x \in [a, b]$ выполняется неравенство $g(x) \leqslant c$. Отсюда имеем (напомним, что $g(x) \geqslant 0$

$$\left|\int_a^b g(x)\,dx - \int_{a+\varepsilon}^{b-\varepsilon} g(x)\,dx\right| = \left|\int_a^{a+\varepsilon} g(x)\,dx + \int_{b-\varepsilon}^{b} g(x)\,dx\right| \leqslant$$

$$\leqslant \int_a^{a+\varepsilon} g(x)\,dx + \int_{b-\varepsilon}^{b} g(x)\,dx \leqslant c\int_a^{a+\varepsilon} dx + c\int_{b-\varepsilon}^{b} dx = 2c\varepsilon,$$

$$0 < \varepsilon < b - a.$$

Из этого неравенства сразу следует равенство (24.51).

В силу неравенства (24.49) и соотношения (24.51), существует такое ε_0, $0 < \varepsilon_0 < b - a$, что $\int\limits_{a+\varepsilon_0}^{b-\varepsilon_0} g(x)\,dx > 0$.

Если бы не существовало точки $\xi \in (a, b)$, в которой $f(\xi) = M$, то непрерывная функция $M - f(x)$ была бы положительной на интервале (a, b), следовательно, и на отрезке $[a + \varepsilon_0, b - \varepsilon_0]$. В частности, она была бы положительной в той точке x_0, в которой она принимает свое наименьшее значение:

$$M - f(x) \geqslant \min_{[a+\varepsilon_0,\, b-\varepsilon_0]} [M - f(x)] = M - f(x_0) > 0,$$

$$x_0 \in [a + \varepsilon_0, b - \varepsilon_0].$$

Поэтому

$$\int_a^b [M - f(x)]g(x)\,dx \geqslant$$

$$\geqslant \int_{a+\varepsilon_0}^{b-\varepsilon_0} [M - f(x)]g(x)\,dx \geqslant [M - f(x_0)] \int_{a+\varepsilon_0}^{b-\varepsilon_0} g(x)\,dx > 0.$$

Последнее противоречит равенству (24.50). Таким образом, в рассматриваемом случае на интервале (a, b) существует такая точка ξ, что $f(\xi) = M = \mu$. □

Следствие теоремы 1 обычно называется *интегральной теоремой о среднем*. Это название обычно объясняется тем, что в нем утверждается существование некоторой точки на отрезке — «средней точки», обладающей определенным свойством, связанным с интегралом от функции.

Формулы (24.37) и (24.38) остаются верными и при $a \geqslant b$.

§ 25.

Определенный интеграл с переменными пределами интегрирования

25.1. Непрерывность интеграла по верхнему пределу интегрирования

Пусть функция $f(x)$ интегрируема на отрезке $[a, b]$. Тогда она интегрируема и на любом отрезке $[a, x]$, где $a \leqslant x \leqslant b$, т. е. для любого $x \in [a, b]$, имеет смысл интеграл $\int\limits_a^x f(t)\,dt$. Рассмотрим функцию

$$F(x) = \int\limits_a^x f(t)\,dt. \tag{25.1}$$

Эта функция F определена на отрезке $[a, b]$ и называется *интегралом с переменным верхним пределом*. Установим ее основные свойства.

ТЕОРЕМА 1. *Если функция f интегрируема на отрезке $[a, b]$, то функция* (25.1) *непрерывна на этом отрезке.*

Доказательство. Пусть $x \in [a, b]$, $x + \Delta x \in [a, b]$. Тогда из формулы (25.1) следует, что

$$F(x + \Delta x) = \int\limits_a^{x + \Delta x} f(t)\,dt =$$
$$= \int\limits_a^x f(t)\,dt + \int\limits_x^{x + \Delta x} f(t)\,dt = F(x) + \int\limits_x^{x + \Delta x} f(t)\,dt,$$

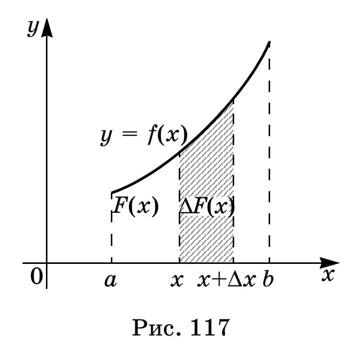


Рис. 117

поэтому (рис. 117)

$$\Delta F = F(x + \Delta x) - F(x) = \int\limits_{x}^{x+\Delta x} f(t)\,dt. \qquad (25.2)$$

Поскольку функция f интегрируема на отрезке $[a, b]$, она ограничена на этом отрезке, т. е. существует такая постоянная $M > 0$, что $|f(x)| \leqslant M$ для всех $x \in [a, b]$. Применяя это неравенство для оценки выражения $|\Delta F|$, получим (см. п. 24.1):

$$|\Delta F| = \left|\int\limits_{x}^{x+\Delta x} f(t)\,dt\right| \leqslant \left|\int\limits_{x}^{x+\Delta x} |f(t)\,dt|\right| \leqslant \left|\int\limits_{x}^{x+\Delta x} M\,dt\right| \leqslant M|\Delta x|.$$

Отсюда следует, что $\lim\limits_{\Delta x \to 0} \Delta F = 0$ для любого $x \in [a, b]$, а это означает непрерывность функции F в каждой точке $x \in [a, b]$. □

25.2. Дифференцируемость интеграла по верхнему пределу интегрирования. Существование первообразной у непрерывной функции

ТЕОРЕМА 2. *Если функция f интегрируема на отрезке $[a, b]$ и непрерывна в точке $x_0 \in [a, b]$, то функция*

$$F(x) = \int\limits_{a}^{x} f(t)\,dt \qquad (25.3)$$

дифференцируема в точке x_0 и

$$F'(x_0) = f(x_0). \qquad (25.4)$$

Доказательство. Покажем, что

$$\lim_{\Delta x \to 0} \frac{\Delta F}{\Delta x} = f(x_0),$$

где

$$\Delta F = F(x_0 + \Delta x) - F(x_0), \quad x_0 + \Delta x \in [a, b].$$

Для этого оценим модуль разности $\frac{\Delta F}{\Delta x} - f(x_0)$. Заметив, что $\frac{1}{\Delta x}\int\limits_{x_0}^{x_0+\Delta x} dt = 1$ и, следовательно,

$$f(x_0) = \frac{1}{\Delta x}\int\limits_{x_0}^{x_0+\Delta x} f(x_0)\,dt,$$

будем иметь

$$\left|\frac{\Delta F}{\Delta x} - f(x_0)\right| = \left|\frac{1}{\Delta x}\int\limits_{x_0}^{x_0+\Delta x} f(t)\,dt - \frac{1}{\Delta x}\int\limits_{x_0}^{x_0+\Delta x} f(x_0)\,dt\right| =$$

$$= \left|\frac{1}{\Delta x}\int\limits_{x_0}^{x_0+\Delta x} [f(t) - f(x_0)]\,dt\right| \leqslant \frac{1}{|\Delta x|}\left|\int\limits_{x_0}^{x_0+\Delta x} |f(t) - f(x_0)|\,dt\right|. \quad (25.5)$$

Пусть задано $\varepsilon > 0$. В силу непрерывности функции f в точке x_0, существует такое $\delta > 0$, что если $|x - x_0| < \delta$ и $x \in [a, b]$, то

$$|f(x) - f(x_0)| < \varepsilon. \quad (25.6)$$

Выберем Δx так, что $|\Delta x| < \delta$. Тогда для значений t на отрезке, по которому ведется интегрирование, будем иметь $|t - x_0| \leqslant |\Delta x| < \delta$ и, следовательно, из неравенств (25.5) и (25.6) получим

$$\left|\frac{\Delta F}{\Delta x} - f(x_0)\right| \leqslant \frac{\varepsilon}{|\Delta x|}\left|\int\limits_{x_0}^{x_0+\Delta x} dt\right| = \varepsilon,$$

а это означает, что существует $\lim\limits_{\Delta x \to 0} \frac{\Delta F}{\Delta x} = f(x_0)$, т. е. существует производная $F'(x_0) = f(x_0)$.

В том случае, когда точка x_0 совпадает с одним из концов отрезка $[a, b]$, под $F'(x_0)$ следует подразумевать соответствующую одностороннюю производную функции $F(x)$. □

Теперь можно решить вопрос о существовании первообразной для произвольной непрерывной функции.

ТЕОРЕМА 3. *Если функция непрерывна на отрезке, то на этом отрезке у нее существует первообразная.*

Доказательство. Если функция f непрерывна на отрезке $[a, b]$, то, по теореме 4 п. 23.6, она и интегрируема на нем. Поэтому, согласно формуле (25.4), справедливой при сделанных предположениях для всех точек $x_0 \in [a, b]$, функция $F(x) = \int\limits_a^x f(t)\,dt$ является первообразной для функции f. □

Таким образом, операция интегрирования с переменным верхним пределом, примененная к непрерывной функции, приводит к первообразной функции, т. е. является операцией, обратной дифференцированию:

$$\frac{d}{dx}\int\limits_a^x f(t)\,dt = f(x), \quad a \leqslant x \leqslant b. \quad (25.7)$$

Это утверждение, называемое *формулой дифференцирования определенного интеграла по верхнему пределу*, является основополагающим для дифференциального и интегрального исчисления. Из него следует, в частности, что любая первообразная функции $f(x)$, непрерывной на отрезке $[a, b]$, имеет вид

$$\int_a^x f(t)\,dt + C, \quad a \leqslant x \leqslant b.$$

Действительно, согласно доказанному, функция $F(x) = \int_a^x f(t)\,dt$ является первообразной для функции $f(x)$, а всякая другая ее первообразная может отличаться от $F(x)$ лишь на постоянную (см. п. 18.1). Таким образом, установлена связь между неопределенным и определенным интегралами, которая имеет вид

$$\int f(x)\,dx = \int_a^x f(t)\,dt + C.$$

Доказанные теоремы показывают, что операция интегрирования с переменным верхним пределом приводит к «улучшению» или «сглаживанию» свойств функции: интегрируемая функция переходит в непрерывную, а непрерывная — в дифференцируемую.

Заметим, что операция дифференцирования в определенном смысле «ухудшает» свойства функции: например, производная непрерывной функции, если она существует, может быть уже разрывной функцией.

Из формулы дифференцирования по верхнему пределу интегрирования, т. е. из формулы (25.7), можно легко получить и формулу дифференцирования по нижнему пределу интегрирования.

Пусть функция f интегрируема на отрезке $[a, b]$. Тогда на этом отрезке определена и функция

$$G(x) = \int_x^b f(t)\,dt, \; a \leqslant x \leqslant b,$$

причем из тождества

$$\int_a^b f(t)\,dt = \int_a^x f(t)\,dt + \int_x^b f(t)\,dt$$

имеем

$$G(x) = \int_a^b f(t)\,dt - F(x). \qquad (25.8)$$

Если функция f непрерывна в точке $x \in [a, b]$, то, как доказано выше, функция F дифференцируема в этой точке. Из формулы (25.8) следует, что в этом случае функция $G(x)$ в точке x также дифференцируема и $\frac{dG(x)}{dx} = -\frac{dF(x)}{dx}$. Таким образом,

$$\frac{d}{dx}\int_x^b f(x)\,dt = -f(t).$$

З а м е ч а н и е. Из формул дифференцирования интеграла от непрерывной функции по верхнему (нижнему) пределу интегрирования следует также, что всякая функция, непрерывная на некотором промежутке (конечном или бесконечном), имеет на нем первообразную. Действительно, пусть, например, функция f непрерывна на интервале (a, b). Выберем произвольную точку $x_0 \in (a, b)$ и положим

$$F(x) = \int_{x_0}^{x} f(t)\,dt.$$

Тогда для всех $x \in (a, b)$ справедливо равенство $F'(x) = f(x)$, т. е. $F(x)$ является первообразной функции $f(x)$ на интервале (a, b).

УПРАЖНЕНИЕ 1. Пусть функция $f(x)$ непрерывна, а $\varphi(x)$ и $\psi(x)$ дифференцируемы всюду в $\boldsymbol{R}$. Доказать следующие обобщения формулы (25.7):

$$\frac{d}{dx}\int_a^{\varphi(x)} f(t)\,dt = f(\varphi(x))\varphi'(x);$$

$$\frac{d}{dx}\int_{\psi(x)}^{\varphi(x)} f(t)\,dt = f(\varphi(x))\varphi'(x) - f(\psi'(x))\psi'(x).$$

25.3. Формула Ньютона—Лейбница

ТЕОРЕМА 4. (основная теорема интегрального исчисления). *Пусть функция f непрерывна на отрезке $[a, b]$. Если функция Φ является произвольной ее первообразной на этом отрезке, то*

$$\int_a^b f(x)\,dx = \Phi(b) - \Phi(a). \qquad (25.9)$$

Эта формула называется *формулой Ньютона—Лейбница.*

Доказательство. Положим $F(x) = \int_a^x f(t)\,dt$. Согласно теореме 3 п. 25.2, функция F является первообразной для функции f на отрезке $[a, b]$. Таким образом, F и Φ — две пер-

вообразные одной и той же функции f на отрезке $[a,\ b]$, поэтому

$$F(x) = \Phi(x) + C,\ \ a \leqslant x \leqslant b,$$

где C — некоторая постоянная, т. е.

$$\int_a^x f(t)\,dt = \Phi(x) + C,\ a \leqslant x \leqslant b.$$

При $x = a$ отсюда следует, что $C = -\Phi(a)$, следовательно,

$$\int_a^x f(t)\,dt = \Phi(x) - \Phi(a).$$

Полагая здесь $x = b$, получим формулу (25.9). □

Для краткости записи часто употребляют обозначение

$$\Phi(x)\Big|_a^b \overset{\text{def}}{=} \Phi(b) - \Phi(a),$$

или

$$[\Phi(x)]_a^b \overset{\text{def}}{=} \Phi(b) - \Phi(a).$$

Заметим, что формула Ньютона—Лейбница (25.9) справедлива и для $a > b$. Действительно, если в ней поменять местами a и b, то ее левая и правая части изменят знак.

Примеры. 1. Найдем $\int_0^1 x^2\,dx$. Известно, что

$$\int x^2\,dx = \frac{x^3}{3} + C,$$

поэтому

$$\int_0^1 x^2\,dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3}.$$

2. Найдем $\int_0^\pi \sin x\,dx$. Имеем

$$\int_0^\pi \sin x\,dx = -\cos x\Big|_2^\pi = -\cos \pi + \cos 0 = 2.$$

25.4*. Существование обобщенной первообразной. Формула Ньютона—Лейбница для обобщенной первообразной

Теорема 3 п. 25.2 и теорема 4 п. 25.3 обобщаются на случай, когда под первообразной понимается обобщенная первообразная (см. п. 18.6*).

ТЕОРЕМА 3*. *Если функция, определенная на отрезке, ограничена на нем и имеет конечное множество точек разрыва, то она имеет на этом отрезке обобщенную первообразную.*

Доказательство. Если функция f ограничена на отрезке $[a, b]$ и имеет конечное множество точек разрыва, то она, согласно критерию интегрируемости Лебега (см. п. 23.10*), интегрируема на этом отрезке, следовательно, имеет смысл функция

$$F(x) = \int_a^x f(x)\,dt,$$

задаваемая этой формулой для всех $x \in [a, b]$. В силу теоремы 1 п. 25.1, функция F непрерывна на отрезке $[a, b]$, а в силу теоремы 2 п. 25.2, для всех точек $x \in [a, b]$, в которых функция f непрерывна (т. е. во всех точках отрезка $[a, b]$, кроме конечного их множества), выполняется условие

$$F'(x) = f(x).$$

Таким образом, функция F является обобщенной первообразной для функции f. □

ТЕОРЕМА 4*. *Пусть функция f, определенная на отрезке, ограничена на нем и множество точек ее разрыва конечное. Если функция Φ является какой-либо обобщенной первообразной функции f на этом отрезке, то справедлива формула Ньютона—Лейбница*

$$\int_a^b f(x)\,dx = \Phi(b) - \Phi(a).$$

Доказательство этой теоремы проводится аналогично доказательству теоремы 4, если только под первообразными понимать обобщенные первообразные.

Покажем теперь, что формула Ньютона—Лейбница имеет место и лишь при предположении существования обобщенной первообразной у интегрируемой функции f, т. е. при условии существования обобщенной первообразной для справедливости формулы Ньютона—Лейбница, не нужно требовать конечности множества точек разрыва функции f (напомним, однако, что конечность множества точек разрыва использовалась при доказательстве *существования* обобщенной первообразной, иначе говоря, конечность мно-

жества точек разрыва ограниченной функции является достаточным условием существования у нее обобщенной первообразной).

ТЕОРЕМА 5. *Пусть функция f интегрируема на отрезке $[a, b]$ и F — ее обобщенная первообразная на этом отрезке. Тогда справедлива формула Ньютона—Лейбница*

$$\int_a^b f(x)\,dx = F(b) - F(a). \qquad (25.10)$$

Доказательство. Согласно определению обобщенной первообразной, функция F непрерывна на отрезке $[a, b]$ и существует такое конечное множество $E_f \subset [a, b]$, что для всех точек $x \in [a, b]\backslash E_f$ выполняется равенство $F'(x) = f(x)$. Обозначим через $a_1, a_2, \ldots, a_m$ точки конечного множества E_f:

$$E_f = \{a_1, a_2, \ldots, a_m\}$$

и рассмотрим какое-либо разбиение $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ отрезка $[a, b]$, содержащее все точки $a_1, \ldots, a_m$. Тогда на каждом отрезке $[x_{i-1}, x_i]$ функция F непрерывна, а внутри него она имеет производную $F'(x) = f(x)$. Поэтому к функции F на указанном отрезке можно применить формулу конечных приращений (теорему Лагранжа о среднем значении):

$$F(x_i) - F(x_{i-1}) = F'(\xi_i)\Delta x_i = f(\xi_i)\Delta x_i, \qquad (25.11)$$

где $\Delta x_i = x_i - x_{i-1}$, $\xi_i \in (x_{i-1}, x_i)$, $i = 1, \ldots, i_\tau$.

Суммируя получившиеся равенства от 1 до i_τ и замечая, что $x_{i_\tau} = b$, $x_0 = a$,

$$\sum_{i=1}^{i_\tau} F(x_i) - F(x_{i-1}) = F(x_{i_\tau}) - F(x_0) = F(b) - F(a),$$

получаем

$$F(b) - F(a) = \sum_{i=1}^{i_\tau} f(\xi_i)\Delta x_i. \qquad (25.12)$$

В правой части этого равенства стоит интегральная сумма Римана функции f.

Пусть $\tau = \tau_n$, $n = 1, 2, \ldots,$ — последовательность разбиений, содержащих точки $a_1, \ldots, a_m$, для которой $|\tau_n| \to 0$ при $n \to \infty$. Переходя к пределу при $n \to \infty$ в (25.12), где

$\tau = \tau_n$, и замечая, что левая часть этого равенства постоянна и равна $F(b) - F(a)$, а правая, в силу интегрируемости функции f, стремится к интегралу $\int\limits_a^b f(x)\,dx$, получим формулу (25.10). □

Из формулы (25.10) следует, что если две интегрируемые функции f и f_1 имеют на отрезке $[a, b]$ одну и ту же обобщенную первообразную F, то интегралы от них по этому отрезку равны, так как они равны числу $F(b) - F(a)$. Это нетрудно доказать и непосредственно, так как в этом случае интегрируемые функции f и f_1 могут отличаться друг от друга только значениями в конечном числе точек.

Замечание. Формула Ньютона—Лейбница иногда записывается в виде

$$\int\limits_a^b F'(x)\,dx = F(b) - F(a). \tag{25.13}$$

Здесь предполагается, что функция F непрерывна на отрезке $[a, b]$ и во всех его точках, кроме некоторого конечного множества, имеет производную F'. Тем самым подынтегральная функция в формуле (25.13) может оказаться определенной не во всех точках отрезка $[a, b]$, и поэтому следует пояснить, что же понимается в этом случае под интегралом $\int\limits_a^b F'(x)\,dx$. В формуле (25.13) дополнительно предполагается, что существует такая интегрируемая на отрезке $[a, b]$ (и тем самым определенная уже в каждой его точке) функция f, для которой F является ее обобщенной первообразной, и, следовательно, существует такое конечное множество E_f, что для всех точек $x \in [a, b] \setminus E_f$ имеет место равенство $F'(x) = f(x)$. Интеграл же $\int\limits_a^b F'(x)\,dx$, по определению, принимается равным интегралу $\int\limits_a^b f(x)\,dx$, т. е.

$$\int\limits_a^b F'(x)\,dx \overset{\text{def}}{=} \int\limits_a^b f(x)\,dx. \tag{25.14}$$

Определение корректно, так как не зависит от выбора указанной функции f: в любом случае она имеет обобщенную первообразную F и, следовательно, в силу теоремы 5, будет получаться одно и то же значение интеграла $\int\limits_a^b f(x)\,dx$,

равное $F(b) - F(a)$. Все сказанное делает естественным следующее определение.

Определение 1. *Функция F, определенная на отрезке $[a, b]$, называется функцией с интегрируемой на этом отрезке производной, если существуют конечное множество $E \subset [a, b]$ и интегрируемая на $[a, b]$ функция f такие, что для любой точки $x \in [a, b] \setminus E$ функция F имеет производную и $F'(x) = f(x)$.*

Иначе говоря, функция F называется *функцией с интегрируемой производной* на некотором отрезке, если на этом отрезке F является обобщенной первообразной интегрируемой функции. Теперь теорему 5 можно перефразировать следующим образом.

ТЕОРЕМА 5*. *Если функция F непрерывна на отрезке $[a, b]$ и имеет на нем интегрируемую производную, то*

$$\int_a^b F'(x)\,dx = F(b) - F(a).$$

УПРАЖНЕНИЕ 2. Доказать, что если непрерывные на отрезке $[a, b]$ функции F_1 и F_2 имеют интегрируемые на этом отрезке производные, то и их произведение F_1F_2 также имеет интегрируемую на $[a, b]$ производную.

§ 26.

Формулы замены переменной в интеграле и интегрирования по частям

26.1. Замена переменной

ТЕОРЕМА 1. *Пусть*:

1) *функция $f(x)$ непрерывна на интервале (a, b)*;

2) *функция $\varphi(t)$ определена и непрерывна вместе со своей производной $\varphi'(t)$ на интервале (α, β), причем для всех $t \in (\alpha, \beta)$ выполняется неравенство $a < \varphi(t) < b$.*

Тогда если $\alpha_0 \in (\alpha, \beta)$, $\beta_0 \in (\alpha, \beta)$, $a_0 = \varphi(a_0)$, $b_0 = \varphi(b_0)$, то

$$\int_{a_0}^{b_0} f(x)\,dx = \int_{\alpha_0}^{\beta_0} f[\varphi(t)]\varphi'(t)\,dx. \qquad (26.1)$$

Эта формула называется формулой *замены переменной* в определенном интеграле или формулой интегрирования *подстановкой*.

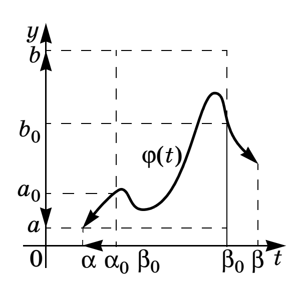


Рис. 118

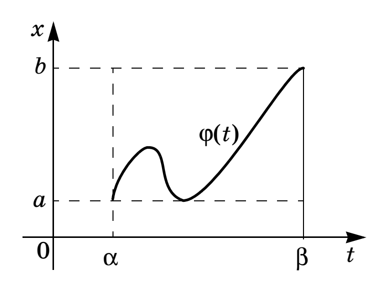


Рис. 119

Доказательство. Прежде всего заметим, что, по условию, функция f заведомо определена на множестве значений функции φ (рис. 118), поэтому имеет смысл сложная функция $f[\varphi(t)]$. В силу сделанных предположений, подынтегральные функции в обеих частях формулы (26.1) непрерывны, поэтому оба интеграла в этой формуле существуют.

Пусть $\Phi(x)$ — какая-либо первообразная функция $f(x)$ на интервале $(a,\ b)$. Тогда для точек t интервала $(\alpha,\ \beta)$ имеет смысл сложная функция $\Phi[\varphi(t)]$, которая является первообразной для функции $f[\varphi(t)]\varphi'(t)$. По формуле Ньютона—Лейбница (см. п. 25.3),

$$\int\limits_{a_0}^{b_0} f(x)\,dx = \Phi(b_0) - \Phi(a_0),$$

$$\int\limits_{\alpha_0}^{\beta_0} f[\varphi(t)]\varphi'(t)\,dt = \Phi[\varphi(\beta_0)] - \Phi[\varphi(\alpha_0)] = \Phi(b_0) - \Phi(a_0).$$

Из этих равенств и следует формула (26.1). □

Как видно из доказательства, формула (26.1) справедлива как при $\alpha_0 \leqslant \beta_0$, так и при $\alpha_0 > \beta_0$.

Интересно отметить, что некоторые значения функции $\varphi(t)$ могут и не принадлежать отрезку $[a_0,\ b_0]$, по которому происходит интегрирование (см. рис. 118) в левой части равенства (26.1).

Если воспользоваться формулой для односторонних производных сложной функции (см. замечание 2 в п. 9.7), то формула (26.1) имеет место для случая, когда f задана на отрезке $[a,\ b]$, функция $\varphi(t)$ — на отрезке $[\alpha,\ \beta]$ и множество значений функции φ содержится в отрезке $[a,\ b]$, причем $a = \varphi(\alpha)$, $b = \varphi(\beta)$ (рис. 119). В этом случае формула замены переменной может быть применена ко всему отрезку $[a,\ b]$:

$$\int\limits_a^b f(x)\,dx = \int\limits_\alpha^\beta f(\varphi(t))]\,\varphi'(t)\,dt. \tag{26.2}$$

Употребляя символ определенного интеграла, мы всегда писали под знаком интеграла выражение $f(x)\,dx$, где x — независимая переменная. При этом, когда давалось определение определенного интеграла, не предполагалось, что $f(x)\,dx$ означает дифференциалы какой-либо функции. Затем (см. п. 25.2) было показано, что по крайней мере для непрерывной функции выражение $f(x)\,dx$ всегда является дифференциалом некоторой функции, а именно ее первообразной $F(x)$: $dF(x) = f(x)\,dx$. Поэтому естественно считать, что в этом случае записи $\int_a^b dF(x)$ и $\int_a^b f(x)\,dx$ равнозначны, т. е.

$$\int_a^b dF(x) = \int_a^b f(x)\,dx.$$

Будем вообще допускать под знаком определенного интеграла любую запись дифференциала, т. е. положим, по определению, для дифференцируемой функции $g(x)$

$$\int_a^b f(x)\,dg(x) = \int_a^b f(x)g'(x)\,dx$$

(если, конечно, интеграл в правой части равенства существует). Например, с помощью этого обозначения формулу (26.2) можно записать в виде

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f[\varphi(t)]\,d\varphi(t).$$

Таким образом, при замене переменного $x = \varphi(t)$ (когда это допустимо) в определенном интеграле $\int_a^b f(x)\,dx$ следует всюду формально заменить x на $\varphi(t)$ и соответственно изменить пределы интегрирования.

Обратим внимание на то, что при применении формулы (26.1) (формулы (26.2)) ее, подобно случаю неопределенного интеграла, можно использовать как слева направо, так и справа налево. Однако в отличие от неопределенного интеграла, где в конце вычисления следует возвращаться к первоначальной переменной интегрирования, здесь этого делать не нужно, так как наша цель найти число, которое, в силу доказанных фор-

мул, равно значению каждого из рассматриваемых интегралов.

Примеры. 1. Вычислим интеграл $\int\limits_0^2 e^{x^2} x dx$. Применив формулу (26.1) справа налево (здесь роль переменной t играет x), получим

$$\int\limits_0^2 e^{x^2} x\, dx = \frac{1}{2}\int\limits_0^2 e^{x^2}\, dx^2 = \frac{1}{2}\int\limits_0^4 e^y\, dy = \frac{1}{2}e^y\Big|_0^4 = \frac{e^4 - 1}{2}.$$

2. Пусть требуется вычислить интеграл $\int\limits_0^{\ln 2}\sqrt{e^x - 1}\, dx$. Попытаемся упростить подынтегральное выражение, положив $\sqrt{e^x - 1} = t$. Иначе говоря, сделаем замену переменного $x = \ln(1 + t^2)$; тогда $dx = \frac{2t\,dt}{1 + t^2}$, и так как при $0 \leqslant t \leqslant 1$ имеем $0 \leqslant x \leqslant \ln 2$, то, применив формулу (26.1) слева направо, получим

$$\int\limits_0^{\ln 2}\sqrt{e^x - 1}\, dx = 2\int\limits_0^1 \frac{t^2 dt}{1 + t^2} =$$

$$= 2\int\limits_0^1\left(1 - \frac{1}{1 + t^2}\right) dt = 2\Big[t - \operatorname{arctg} t\Big]_0^1 = \frac{4 - \pi}{2}.$$

УПРАЖНЕНИЕ 1. Доказать, что если функция f непрерывна на отрезке $[a, b]$ и для всех $t \in [0, b - a]$ имеет место равенство $f(a + t) = f(b - t)$, то

$$\int\limits_a^b x f(x)\, dx = \frac{a + b}{2}\int\limits_a^b f(x)\, dx.$$

Заметим, что в случае *непрерывных* функций формулы (24.31) для четных функций и (24.32) для периодических функций (см. примеры в п. 24.1) сразу следуют из теоремы 1 о замене переменного в интеграле.

В самом деле, если функция f непрерывна на отрезке $[a, b]$, $0 \leqslant a < b$, а $f^*(x) = f(-x)$, $-b \leqslant x \leqslant -a$, то равенство

$$\int\limits_{-b}^{-a} f^*(x)\, dx = \int\limits_a^b f(x)\, dx$$

сразу получится, если в интеграле $\int\limits_{-b}^{-a} f^*(x)\,dx$ сделать замену переменного $x = t - a - b$.

Если же функция f непрерывна на числовой оси и имеет период $T > 0$, то для доказательства формулы

$$\int\limits_{a}^{b+T} f(x)\,dx = \int\limits_{b}^{b+T} f(x)\,dx,$$

имеющей место для любых чисел a и b, выберем, например, при $b \geqslant a$ такое неотрицательное целое n (см. пример 2 в п. 24.1), что $a + nT \leqslant b < a + (n+1)T$, и представим интеграл $\int\limits_{a}^{a+T} f(x)\,dx$ в виде суммы двух интегралов следующим образом:

$$\int\limits_{a}^{a+T} f(x)\,dx = \int\limits_{a}^{b+nT} f(x)\,dx + \int\limits_{b-nT}^{a+T} f(x)\,dx.$$

Сделав в первом интеграле правой части равенства замену переменного $x = t - (n+1)T$, а во втором — $x = t - nT$, получим

$$\int\limits_{a}^{a+T} f(x)\,dx = \int\limits_{a+(n+1)T}^{b+T} f(t)\,dt + \int\limits_{b}^{a+(n+1)T} f(t)\,dt = \int\limits_{b}^{b+T} (t)\,dt.$$

Формула замены переменных в определенном интеграле с помощью формулы Ньютона—Лейбница может быть обобщена и на случай, когда функция f интегрируема на отрезке и имеет на нем первообразную (см. теорему 5^*).

26.2. Интегрирование по частям

Т Е О Р Е М А 2. *Если функции* $u(x)$ *и* $v(x)$ *непрерывно дифференцируемы на отрезке* $[a, b]$, *то*

$$\int\limits_{a}^{b} u\,dv = uv\Big|_{a}^{b} - \int\limits_{a}^{b} v\,du. \qquad (26.3)$$

Эта формула называется *формулой интегрирования по частям* для определенного интеграла.

Д о к а з а т е л ь с т в о. Имеем

$$\int\limits_{a}^{b} (uv)'\,dx = \int\limits_{a}^{b} (uv' + u'v)\,dx = \int\limits_{a}^{b} u\,dv + \int\limits_{a}^{b} v\,du. \qquad (26.4)$$

Все написанные интегралы существуют, ибо подынтегральные функции непрерывны. Согласно формуле Ньютона—Лейбница (25.11), имеем

$$\int_a^b (uv)'\,dx = |uv|_a^b. \qquad (26.5)$$

Сравнив формулы (26.4) и (26.5), получим равенство

$$\int_a^b u\,du + \int_a^b v\,du = |uv|_a^b, \qquad (26.6)$$

откуда и следует формула (26.3). □

Теорема 2 легко обобщается на случай кусочно-непрерывно дифференцируемых функций. Определим эти функции.

Пусть функция $f(x)$ определена на отрезке $[a, b]$, существует такое разбиение $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ отрезка $[a, b]$, что $f(x)$ непрерывна на каждом интервале (x_{i-1}, x_i), и существуют конечные пределы $f(x_{i-1}+0)$, $f(x_i-0)$, $i = 1, 2, \dots, i_\tau$. (Следовательно, функция f кусочно-непрерывна на отрезке $[a, b]$, см. определение 6 в п. 23.10*.)

$$f_i(x) = \begin{cases} f(x), & \text{если } x_{i-1} < x < x_i, \\ f(x_{i-1}+0), & \text{если } x = x_{i-1}, \\ f(x_i-0), & \text{если } x = x_i. \end{cases}$$

Определение 1. *Если каждая функция $f_i(x)$, $i = 1, 2, \dots, k$, (непрерывно) дифференцируема на отрезке $[x_{i-1}, x_i]$, то функция $f(x)$ называется кусочно (непрерывно) дифференцируемой на отрезке $[a, b]$.*

ТЕОРЕМА 2*. *Пусть функция $u(x)$ и функция $v(x)$ кусочно непрерывно дифференцируемы на отрезке $[a, b]$; тогда для них справедлива формула* (26.3) *интегрирования по частям.*

Доказательство теоремы 2 остается в силе и в этом случае. Действительно, произведение функций u, v и его производная $(uv)' = uv' + u'v$ кусочно-непрерывны. Поэтому, согласно теореме 4* п. 25.4*, к интегралу в левой части равенства (26.5) можно также применить формулу Ньютона—Лейбница. □

Примеры. 1. Найдем значение интеграла $\int_1^2 \ln x\, dx$. Применим формулу интегрирования по частям:

$$\int_1^2 \ln x\, dx = x\ln x\Big|_1^2 - \int_1^2 dx = 2\ln 2 - 1.$$

2. Покажем, что для любого $n = 1, 2, \ldots$

$$I_n = \int_0^{\pi/2} \sin^n x\, dx = \int_0^{\pi/2} \cos^n x\, dx =$$

$$= \begin{cases} \dfrac{(n-1)!!}{n!!}\dfrac{\pi}{2} & \text{при } n \text{ четном,} \\ \dfrac{(n-1)!!}{n!!} & \text{при } n \text{ нечетном.} \end{cases} \qquad (26.7)$$

Под $n!!$, $n \in N$, $a \geqslant 1$, понимается произведение всех натуральных чисел, не превосходящих n и обладающих той же четностью, что и число n, и по определению $0!! = 1$.

Равенство интегралов, входящих в (26.7), легко установить с помощью замены переменного $x = \pi/2 - t$.

Положим $I_0 = \int_0^{\pi/2} dx = \pi/2$. Интегрируя по частям, имеем

$$I_n = \int_0^{\pi/2} \sin^n x\, dx = \int_0^{\pi/2} \sin^{n-1} x\, d(-\cos x) =$$

$$= -\sin^{n-1} x \cos x\Big|_0^{\pi/2} + (n-1)\int_0^{\pi/2} \sin^{n-2} x \cos^2 x\, dx =$$

$$= (n-1)\int_0^{\pi/2} \sin^{n-2} x(1 - \sin^2 x)\, dx = (n-1)I_{n-2} - (n-1)I_n,$$

отсюда

$$I_n = \frac{n-1}{n} I_{n-2}.$$

Заметим, что $I_0 = \pi/2$, $I_1 = \int_0^{\pi/2} \sin x\, dx = 1$. Поэтому при $n = 2k + 1$, т. е. нечетном, будем иметь

$$I_{2k+1} = \frac{2k}{2k+1} I_{2k-1} = \ldots = \frac{2k(2k-2)\ldots 2}{(2k+1)(2k-1)\ldots 1} I_1 = \frac{2k!!}{(2k+1)!!},$$

а при $n = 2k$, т. е. четном, —

$$I_{2k} = \frac{2k-1}{2k} I_{2k-2} = \ldots = \frac{(2k-1)(2k-3)\ldots 1}{2k(2k-2)\ldots 2} I_0 = \frac{(2k-1)!!}{(2k)!!}\frac{\pi}{2},$$

$$k = 1, 2, \ldots . \ \square$$

Из формулы (26.7) легко получается ф о р м у л а В а л л и с а[1], которая понадобится в дальнейшем:

$$\frac{\pi}{2} = \lim_{n \to \infty} \frac{1}{2n+1}\left[\frac{(2n)!!}{(2n-1)!!}\right]^2. \tag{26.8}$$

Докажем ее. Интегрируя неравенство

$$\sin^{2n+1} x \leqslant \sin^{2n} x \leqslant \sin^{2n-1} x,\ 0 \leqslant x \leqslant \frac{\pi}{2},$$

по отрезку $[0, \pi/2]$, будем иметь

$$\int_0^{\pi/2} \sin^{2n+1}x\,dx \leqslant \int_0^{\pi/2} \sin^{2n}x\,dx \leqslant \int_0^{\pi/2} \sin^{2n-1}x\,dx$$

(нетрудно показать, что в действительности, в силу свойства 9^0 определенных интегралов (см. п. 24.1), здесь имеют место строгие неравенства). Согласно (26.7),

$$\frac{(2n)!!}{(2n+1)!!} \leqslant \frac{(2n-1)!!}{(2n)!!}\frac{\pi}{2} \leqslant \frac{(2n-2)!!}{(2n-1)!!},$$

откуда

$$x_n \overset{\text{def}}{=} \frac{1}{2n+1}\left[\frac{(2n)!!}{(2n-1)!!}\right]^2 \leqslant \frac{\pi}{2} \leqslant \frac{1}{2n}\left[\frac{(2n)!!}{(2n-1)!!}\right]^2 \overset{\text{def}}{=} y_n. \tag{26.9}$$

В силу этого неравенства, при $n \to \infty$

$$y_n - x_n = \frac{1}{2n}\frac{1}{2n+1}\left[\frac{2n!!}{(2n-1)!!}\right]^2 \leqslant \frac{1}{2n}\frac{\pi}{2} \to 0,$$

поэтому $\lim\limits_{n \to \infty}(y_n - x_n) = 0$, т. е. длины отрезков $[x_n, y_n]$, содержащих $\frac{\pi}{2}$, стремятся к нулю и, следовательно, $\lim\limits_{n \to \infty} x_n = \frac{\pi}{2}$, $\lim\limits_{n \to \infty} y_n = \frac{\pi}{2}$. Первое из этих равенств, согласно определению x_n (см. (26.9)), и означает справедливость формулы Валлиса.

26.3*. Вторая теорема о среднем значении для определенного интеграла

ЛЕММА 1. *Пусть f — непрерывная, а g — возрастающая неотрицательная непрерывно дифференцируемая на от-*

[1] Дж. Валлис (1616—1703) — английский математик.

резке $[a, b]$ *функция. Тогда существует такая точка* $\xi \in$ $\in [a, b]$, *что*

$$\int_a^b g(x)f(x)\,dx = g(b)\int_\xi^b f(x)\,dx. \qquad (26.10)$$

Доказательство. Рассмотрим функцию

$$F(x) \stackrel{\text{def}}{=} \int_x^b f(t)\,dt, \quad a \leqslant x \leqslant b. \qquad (26.11)$$

Функция F, являясь интегралом с переменным нижним пределом интегрирования от интегрируемой (даже непрерывной) функции f, непрерывна на отрезке $[a, b]$ и поэтому достигает на нем своего наибольшего и наименьшего значения. Если

$$m = \min_{[a,b]} F(x), \quad M = \max_{[a,b]} F(x), \qquad (26.12)$$

то, очевидно,

$$m \leqslant F(x) \leqslant M, \; x \in [a, b]. \qquad (26.13)$$

Заметив, что $dF(x) = -f(x)\,dx$ и проинтегрировав по частям интеграл в левой части равенства (26.10), получим

$$\int_a^b g(x)f(x)\,dx = -\int_a^b g(x)dF(x) = -g(x)F(x)\Big|_a^b + \int_a^b F(x)g'(x)\,dx =$$
$$= g(a)F(a) + \int_a^b F(x)g'(x)\,dx, \qquad (26.14)$$

так как, в силу (26.11), $F(b) = 0$.

Функция g возрастающая, поэтому имеем $g'(x) \geqslant 0$ для всех $x \in [a, b]$. Применив это неравенство, неравенства (26.13) и заметив, что из неотрицательности g на $[a, b]$ следует, в частности, что $g(a) \geqslant 0$, получим оценки

$$g(a)F(a) + \int_a^b F(x)g'(x)\,dx \leqslant Mg(a) + M\int_a^b g'(x)\,dx =$$
$$= Mg(a) + M[g(b) - g(a)] = Mg(b),$$

$$g(a)F(a) + \int_a^b F(x)g'(x)\,dx \geqslant mg(a) + m[g(b) - g(a)] = mg(b).$$

Таким образом (см. (26.14)), имеем

$$mg(b) \leqslant \int_a^b g(x)f(x)\,dx \leqslant Mg(b).$$

Если $g(b) = 0$, то из неотрицательности и возрастания функции g следует, что $g(x) \equiv 0$ на $[a, b]$. В этом случае формула (26.10) справедлива при любом выборе $\xi \in [a, b]$.

Если же $g(b) > 0$, то

$$m \leqslant \frac{1}{g(b)}\int_a^b g(x)f(x)\,dx \leqslant M.$$

Непрерывная на отрезке $[a, b]$ функция F принимает на этом отрезке любое значение, лежащее между ее минимальным значением m и максимальным M (см. (26.12)), поэтому существует такая точка $\xi \in [a, b]$, что

$$F(\xi) = \frac{1}{g(b)}\int_a^b g(x)f(x)\,dx.$$

В силу определения (26.11), это и есть формула (26.10). □

ТЕОРЕМА 3 (теорема Бонне[1]**).** *Пусть f — непрерывная, а g — монотонная непрерывно дифференцируемая на отрезке* $[a, b]$ *функция. Тогда существует такая точка* $\xi \in [a, b]$, *что*

$$\int_a^b g(x)f(x)\,dx = g(a)\int_a^\xi f(x)\,dx + g(b)\int_\xi^b f(x)\,dx. \qquad (26.15)$$

Доказательство. Допустим сначала, что функция g возрастает на отрезке $[a, b]$; тогда функция $h(x) \overset{\text{def}}{=} g(x) - g(a)$, $a \leqslant x \leqslant b$, будет неотрицательной возрастающей непрерывно дифференцируемой на отрезке $[a, b]$ функцией. Поэтому, согласно лемме, существует такое $\xi \in [a, b]$, что

$$\int_a^b h(x)f(x)\,dx = h(b)\int_\xi^b f(x)\,dx.$$

Подставив сюда выражение для $h(x)$, имеем

$$\int_a^b [g(x) - g(a)]f(x)\,dx = [g(b) - g(a)]\int_\xi^b f(x)\,dx,$$

откуда

$$\int_a^b g(x)f(x)\,dx =$$

$$= g(a)\int_a^b f(x)\,dx - g(a)\int_\xi^b f(x)\,dx + g(b)\int_\xi^b f(x)\,dx =$$

$$= g(a)\int_a^\xi f(x)\,dx + g(b)\int_\xi^b f(x)\,dx,$$

т. е. получилась формула (26.15).

Если функция g убывает на отрезке $[a, b]$, то для доказательства теоремы достаточно применить формулу (26.15) к функции $-g$, которая, очевидно, возрастающая. □

[1] О. Бонне (1819—1892) — французский математик.

Отметим, что теорема 2 справедлива и при более слабых ограничениях: от функции f достаточно потребовать лишь ее интегрируемости, а от g — ее монотонности.

26.4. Интегралы от векторных функций

Аналогично тому, как были определены интегралы от числовых функций, можно определить и интегралы от векторных функций, значения которых принадлежат n-мерному векторному пространству $\boldsymbol{R}^n$ (см. п. 14.4).

Пусть $\boldsymbol{r}(t) \in \boldsymbol{R}^n$, $a \leqslant t \leqslant b$, — векторная функция, $\tau = \{t_i\}_{i=0}^{i=i_\tau}$ — разбиение отрезка $[a, b]$, $\xi_i \in [t_{i-1}, t_i]$, $\Delta t_i = t_i - t_{i-1}$, $i = 1, 2, \ldots, i_\tau$, $|\tau|$ — мелкость разбиения τ. Если при любом указанном выборе точек ξ_i существует предел[1] $\lim\limits_{|\tau| \to 0} \sum\limits_{i=1}^{i_\tau} \boldsymbol{r}(\xi_i)\Delta t_i$, не зависящий от выбора последовательности разбиений, то он называется *интегралом от функции* $\boldsymbol{r}(t)$ *по отрезку* $[a, b]$ и обозначается $\int\limits_a^b \boldsymbol{r}(t)\, dx$. При фиксированных a и b он представляет собой фиксированный вектор в $\boldsymbol{R}^n$.

Пусть $\boldsymbol{r}(t) = (x_1(t), \ldots, x_n(t))$. При сложении векторов складываются их координаты, при умножении векторов на число их координаты умножаются на то же число, а предел векторной функции равен вектору, координаты которого являются пределами ее соответствующих координат, поэтому

$$\int\limits_a^b \boldsymbol{r}(t)\, dt = \left(\int\limits_a^b x_1(t)\, dt, \ldots, \int\limits_a^b x_n(t)\, dt \right).$$

В силу этого равенства, многие свойства интегралов от числовых функций переносятся на интегралы от векторных функций. В частности, векторная функция $\boldsymbol{F}(t)$, определенная на некотором конечном или бесконечном промежутке Δ числовой прямой, называется *первообразной для*

[1] Понятие предела в этом случае определяется либо с помощью предела векторной последовательности, либо эквивалентным образом на (ε—δ)-языке аналогично случаю скалярных функций, рассмотренному в п. 23.1.

данной функции $\boldsymbol{r}(t) \in \boldsymbol{R}^n$, определенной на том же промежутке Δ, если во всех его точках t имеет место равенство $\boldsymbol{F}'(t) = \boldsymbol{r}(t)$.

Для векторных функций справедливо предложение, аналогичное основной теореме интегрального исчисления (см. теорему 4 п. 25.3):

если векторная функция $\boldsymbol{r}(t) \in \boldsymbol{R}^n$ *непрерывна на отрезке* $[a, b]$, *то у нее существует на этом отрезке первообразная и для любой первообразной* $\boldsymbol{F}(t)$ *функции* $\boldsymbol{r}(t)$ *справедлива формула*

$$\int_a^b \boldsymbol{r}(t)\, dt = \boldsymbol{F}(b) - \boldsymbol{F}(a),$$

называемая, как и в случае скалярных функций, формулой Ньютона—Лейбница.

Справедливость этого утверждения следует из справедливости формулы Ньютона—Лейбница для всех координат функции $\boldsymbol{r}(t)$.

З а м е ч а н и е. В п. 15.2 была доказана следующая теорема: если векторная функция $\boldsymbol{r}(t)$ непрерывна на отрезке $[a, b]$ и дифференцируема внутри него, то существует такая точка $\xi \in (a, b)$, что

$$|\boldsymbol{r}(b) - \boldsymbol{r}(a)| \leqslant |\boldsymbol{r}'(\xi)|(b - a).$$

Приведенное в п. 15.2 доказательство этого утверждения имело искусственный характер — надо было догадаться воспользоваться некоторой вспомогательной функцией. С помощью понятия интеграла (предполагая непрерывность производной рассматриваемой векторной функции) доказательство можно провести более естественно.

Пусть векторная функция $\boldsymbol{r}(t) \in \boldsymbol{R}^n$ имеет непрерывную на отрезке $[a, b]$ производную. Тогда, применяя формулу Ньютона—Лейбница, имеем

$$|\boldsymbol{r}(b) - \boldsymbol{r}(a)| = \left|\int_a^b \boldsymbol{r}'(t)\, dt\right| \leqslant \int_a^b |\boldsymbol{r}'(t)|\, dt.$$

В правой части получился интеграл от непрерывной скалярной функции. Согласно интегральной теореме о среднем (см. следствие из теоремы 1 в п. 24.2), существует такая точка $\xi \in (a, b)$, что $\int_a^b |\boldsymbol{r}'(t)|\, dt = |\boldsymbol{r}'(\xi)|(b - a)$; следовательно,

$$|\boldsymbol{r}(b) - \boldsymbol{r}(a)| \leqslant |\boldsymbol{r}'(\xi)|(b - a), \ \ \xi \in (a, b). \ \square$$

§ 27.

Мера плоских открытых множеств

27.1. Определение меры (площади) открытого множества

Для того чтобы выяснить геометрический смысл определенного интеграла, рассмотрим понятие площади множеств, лежащих на плоскости (их называют плоскими множествами). Для наших целей достаточно ограничиться рассмотрением одного класса множеств, состоящего из так называемых открытых множеств. Для более широкого класса множеств понятия площади и объема рассмотрены далее (см. п. 44.1).

Пусть задана плоскость $\boldsymbol{R}^2$. Нам удобно ввести понятие *открытого круга* $U(P_0; \varepsilon)$ с центром в точке $P_0 \in \boldsymbol{R}^2$ и радиуса $\varepsilon > 0$ как множества всех точек P плоскости $\boldsymbol{R}^2$, отстоящих от точки P_0 на расстоянии, меньшем ε:

$$U(P_0; \varepsilon) = \{P \in \boldsymbol{R}^2 : |PP_0| < \varepsilon\},$$

где $|PP_0|$ — длина отрезка с концами в точках P_0 и P.

Открытый круг $U(P_0; \varepsilon)$ называется также ε-*окрестностью точки* P_0.

Множество $\{P \in \boldsymbol{R}^2 : |PP_0| \leqslant \varepsilon\}$, в котором строгое неравенство $|PP_0| < \varepsilon$ заменено на нестрогое $|PP_0| \leqslant \varepsilon$, обычно называют *замкнутым кругом*.

Точка множества, лежащего на плоскости, называется его *внутренней точкой*, если у нее существует ε-окрестность, содержащаяся в этом множестве.

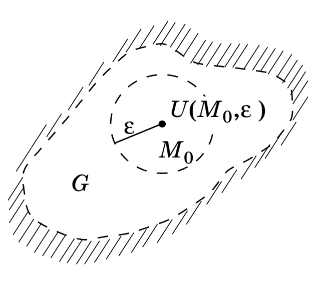


Рис. 120

Множество, все точки которого являются его внутренними точками, называется *открытым множеством*.

В символичной записи это определение выглядит следующим образом: G — открытое множество $\overset{\text{def}}{\Leftrightarrow}$ $\overset{\text{def}}{\Leftrightarrow}$ $\forall M \in G \ \exists\ \varepsilon > 0$: $U(M; \varepsilon) \subset G$ (рис. 120).

Легко видеть, что множество внутренних точек любого множества, лежащего на плоскости, является открытым множеством на этой плоскости. Пустое множество по определению считается открытым.

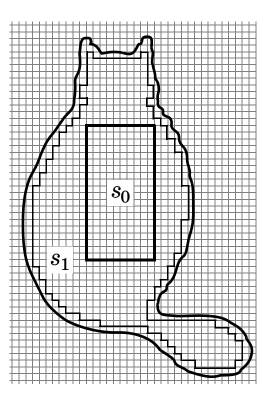


Рис. 121

Рассмотрим плоскость, на которой зафиксирована некоторая прямоугольная система координат. Обозначим через T_0 разбиение этой плоскости на квадраты, получающиеся при проведении всевозможных прямых $x = p$, $y = q$, $p = 0, \pm1, \pm2, \ldots$, $q = 0, \pm1, \pm2, \ldots$, т. е. квадраты вида $\{(x, y): p \leqslant x \leqslant p + 1, q \leqslant y \leqslant q + 1\}$.

Такое разбиение назовем *квадрильяжем плоскости ранга* 0, а указанные квадраты — *квадратами нулевого ранга*. Разобьем каждый из квадратов нулевого ранга на 100 равных квадратов прямыми, параллельными осям координат (любые две соседние параллельные прямые отстоят друг от друга на расстояние 1/10). Совокупность получившихся квадратов обозначим T_1. Продолжая этот процесс дальше, получим квадрильяжи T_m, $m = 1, 2, \ldots$, плоскости, состоящие из квадратов, образовавшихся в результате проведения всевозможных прямых вида

$$x = p/10^m,\ y = q/10^m,\ p = 0, \pm1, \pm2, \ldots,\ q = 0, \pm1, \pm2, \ldots,$$

и, следовательно, со сторонами длины 10^{-m}. Квадраты, принадлежащие квадрильяжу T_m, т. е. квадраты вида $\left\{(x, y): \frac{p}{10^m} \leqslant x \leqslant \frac{p+1}{10^m},\ \frac{q}{10^m} \leqslant y \leqslant \frac{q+1}{10^m}\right\}$, будем называть *квадратами ранга* m, $m = 0, 1, 2, \ldots$.

Пусть G — плоское открытое множество. Обозначим через $s_0 = s_0(G)$ совокупность точек всех квадратов нулевого ранга, лежащих во множестве G, а через $s_1 = s_1(G)$ — совокупность точек всех квадратов первого ранга, лежащих в G. Вообще через $s_m = s_m(G)$ обозначим совокупность всех квадратов ранга m, лежащих в множестве G, $m = 1, 2, \ldots$. Очевидно (рис. 121), что

$$s_0 \subset s_1 \subset \ldots \subset s_m \subset \ldots \subset G. \tag{27.1}$$

Множества s_0, s_1, ... , s_m, ... представляют собой «многоугольники», составленные из конечного или бесконечного числа квадратов соответствующего ранга. В случае, если s_m состоит из конечного числа квадратов, обозначим площадь многоугольника s_m через пл. s_m, если же s_m состоит из бесконечного числа квадратов, положим пл. $s_m = +\infty$. Если какое-то s_{m_0} состоит из бесконечного числа квадратов, то и все следующие s_m, $m \geqslant m_0$, также состоят из бесконечного числа квадратов.

Из включений (27.1), в силу соглашения об использовании символа $+\infty$ (см. п. 3.1), следует, что всегда

$$\text{пл. } s_0 \leqslant \text{пл. } s_1 \leqslant \ldots \leqslant \text{пл. } s_m \leqslant \ldots . \qquad (27.2)$$

Возможны два случая.

1. Все пл. s_m конечны; тогда (27.2) является монотонно возрастающей числовой последовательностью, и поэтому она имеет либо конечный предел, либо стремится к $+\infty$. Этот предел в данном случае и называется *площадью*, или *мерой*, *открытого множества G* и обозначается mes G[1].

2. Если же существует такой номер m_0, что пл. $s_{m_0} = +\infty$, то пл. $s_m = +\infty$ и для всех номеров $m \geqslant m_0$. В этом случае положим mes $G = +\infty$.

Согласно определению предела последовательности элементов расширенной числовой прямой $\overline{\boldsymbol{R}}$ (см. п. 4.2), последовательность элементов a_n, $n = 1, 2, \ldots$, принадлежащих расширенному множеству действительных чисел $\boldsymbol{R}$ таких, что начиная с некоторого номера они все равны $+\infty$: $\lim\limits_{n \to \infty} a_n = +\infty$. Используя это понятие, оба рассмотренных выше случая можно объединить. Сформулируем окончательное определение.

Определение 1. *Предел* $\lim\limits_{m \to \infty}$ *пл.* $s_m(G)$ (*конечный или бесконечный*) *называется площадью, или мерой, открытого множества G и обозначается* mes G:

$$\text{mes } G = \lim_{m \to \infty} \text{пл. } s_m(G). \qquad (27.3)$$

[1] От франц. mésure — мера, размер.

Такое определение меры открытого множества естественно, так как последовательность множеств s_m, $m = 1, 2, \ldots$, исчерпывает открытое множество, т. е. $\bigcup\limits_{m=1}^{\infty} s_m = G$, иначе говоря, для любой точки $P \in G$ существует такой многоугольник s_{m_0}, что $P \in s_{m_0}$.

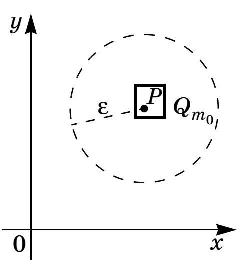


Рис. 122

Действительно, какова бы ни была точка $P \in G$, в силу открытости множества G, существует сферическая окрестность $U(P; \varepsilon) \subset G$, $\varepsilon > 0$. Заметив теперь, что диаметр квадрата ранга m равен $10^{-m}\sqrt{2}$, выберем m_0 так, чтобы

$$\frac{1}{10^{m_0}} < \frac{\varepsilon}{\sqrt{2}}. \tag{27.4}$$

Для всякой точки плоскости существует по крайней мере один квадрат каждого ранга, содержащий эту точку. Пусть Q_{m_0} — квадрат ранга m_0, содержащий точку P. В силу неравенства (27.4) $Q_{m_0} \subset U(P; \varepsilon)$, значит, $Q_{m_0} \subset G$ и, следовательно, $Q_{m_0} \subset s_{m_0}$, но $P \in Q_{m_0}$, поэтому $P \in s_{m_0}$ (рис. 122). □

Если открытое множество G ограничено, то всегда $\text{mes}\, G < +\infty$. В самом деле, если G ограничено, то существует замкнутый квадрат Q, содержащий множество G ($G \subset Q$) и являющийся объединением квадратов нулевого ранга; тогда $s_m(G) \subset Q$ при любом $m = 0, 1, \ldots$, и значит, пл. $s_m(G) \leqslant$ $\leqslant$ пл. Q.

Таким образом, последовательность (27.2) ограничена сверху и, значит, предел (27.3) конечен.

Задача 17. Доказать, что мера плоского открытого множества не зависит от выбора прямоугольной системы координат на плоскости, на которой оно расположено.

Пусть открытое плоское множество S является многоугольником, а $\overline{S}$ — многоугольник, получающийся из многоугольника S добавлением к нему всех точек ограничивающей его ломаной. Многоугольник S называется открытым, а $\overline{S}$ — замкнутым[1]. Из курса элементарной математики известно,

---

[1] См. также п. 44.2 (квадрируемые множества).

что площади многоугольников S и $\overline{S}$ равны и совпадают с определенной нами мерой открытого множества S:

$$\text{пл. } \overline{S} = \text{пл. } S = \text{mes } S.$$

Аналогичное утверждение имеет место и для «открытых и замкнутых круговых секторов».

27.2. Свойства меры открытых множеств

ТЕОРЕМА 1 (монотонность меры). *Если G и Γ — плоские открытые множества и*

$$G \subset \Gamma, \tag{27.5}$$

то

$$\text{mes } G \leqslant \text{mes } \Gamma. \tag{27.6}$$

Доказательство. Обозначим, как и выше, через $s_m(G)$ и $s_m(\Gamma)$ совокупности квадратов ранга m, лежащих соответственно в множествах G и Γ, $m = 1, 2, \ldots$. Тогда из условия (27.5) следует, что

$$s_m(G) \subset s_m\Gamma,$$

откуда

$$\text{пл. } s_m(G) \leqslant \text{пл. } s_m\Gamma. \tag{27.7}$$

В том случае, когда оба множества $s_m\{G\}$ и $s_m(\Gamma)$ состоят из конечного числа квадратов, это следует из того, что площадь объемлющего многоугольника не меньше площади объемлемого, а в том случае, когда хоть одно из множеств $s_m(G)$ и $s_m(\Gamma)$ содержит бесконечно много квадратов, — из соглашения об употреблении символа $+\infty$.

Переходя к пределу в неравенстве (27.7) при $m \to \infty$, в силу (27.3), получим неравенство (27.6). □

ТЕОРЕМА 2. *Пусть G и G_k, $k = 1, 2, \ldots$, — плоские открытые множества, $G_1 \subset G_2 \subset \ldots G_k \subset \ldots$ и $G = \bigcup\limits_{k=1}^{\infty} G_k$, тогда*

$$\lim_{k \to \infty} \text{mes } G_k = \text{mes } G. \tag{27.8}$$

Заметим, что если при некотором k_0 имеет место $\text{mes } G_k = +\infty$, то, согласно теореме 1, и для всех $k \geqslant k_0$ также

mes $G_k = +\infty$; в этом случае равенство (27.8) означает, что mes $G = +\infty$.

Докажем предварительно лемму.

ЛЕММА. *Если G_k, $k = 1, 2, \ldots$, — открытые плоские множества*

$$G_1 \subset G_2 \subset \ldots \subset G_k \subset G_{k+1} \subset \ldots, \ G = \bigcup_{k=1}^{\infty} G_k, \qquad (27.9)$$

s — множество, состоящее из конечного множества квадратов некоторого ранга m, и $s \subset G$, то существует такой номер k, что

$$s \subset G_k.$$

Доказательство леммы. Пусть номера k, для которого выполняется утверждение леммы, не существует. Это означает, что для любого номера $k = 1, 2, \ldots$ найдется такая точка M_k, что $M_k \in s$, но $M_k \notin G_k$.

По условию множество s состоит из конечного множества квадратов ранга m. Поэтому среди этих квадратов найдется по крайней мере один квадрат

$$Q = \left\{(x; y) : a \leqslant x \leqslant a + \frac{1}{10^m}; b \leqslant y \leqslant b + \frac{1}{10^m}\right\},$$

содержащий бесконечно много точек $M_k = (x_k, y_k)$, следовательно, и некоторую подпоследовательность последовательности $\{M_k\}$.

Абсциссы и ординаты точек этой подпоследовательности принадлежат соответственно отрезкам $[a,\ a + 1/10^m]$ и $[b,\ b + 1/10^m]$. Применив теорему Больцано—Вейерштрасса (см. п. 4.6), сначала выделим из абсцисс точек рассматриваемой подпоследовательности сходящуюся подпоследовательность, а затем из подпоследовательности ординат с теми же номерами, что и у полученной сходящейся подпоследовательности абсцисс, выделим снова сходящуюся подпоследовательность.

В результате получим такую подпоследовательность $M_{k_n} = (x_{k_n}, y_{k_n}) \in Q$, $n = 1, 2, \ldots$, что числовые подпоследовательности $\{x_{k_n}\}$ и $\{y_{k_n}\}$ сходятся.

Пусть

$$\lim_{n \to \infty} x_{k_n} = \xi, \quad \lim_{n \to \infty} y_{k_n} = \eta, \qquad (27.10)$$

тогда

$$M_0 \overset{\text{def}}{=} (\xi, \eta) \in Q \subset s \subset G = \bigcup_{k=1}^{\infty} G_k,$$

поэтому найдется такой номер k_0, что $M_0 \in G_{k_0}$. Множество G_{k_0} открытое. Следовательно, существует такое $\varepsilon > 0$, что ε-окрестность $U(M_0; \varepsilon)$ точки M_0 содержится в G_{k_0}:

$$U(M_0; \varepsilon) \subset G_{k_0}. \tag{27.11}$$

Для этого $\varepsilon > 0$ в силу условий (27.10) найдется такой номер n_0, что

$$k_{n_0} > k_0, \tag{27.12}$$

и

$$|x_{k_{n_0}} - \xi| < \frac{\varepsilon}{\sqrt{2}}, \ |y_{k_{n_0}} - \eta| < \frac{\varepsilon}{\sqrt{2}}$$

и, следовательно, для точки $M_{k_{n_0}} = (x_{k_{n_0}}, y_{k_{n_0}})$ выполняется неравенство

$$|M_0 M_{k_{n_0}}| = \sqrt{\left(x_{k_{n_0}} - \xi\right)^2 + (y_{n_{k_0}} - \eta)^2} < \varepsilon,$$

поэтому

$$M_{k_{n_0}} \in U(M_0; \varepsilon) \underset{(27.11)}{\subset} G_{k_1} \underset{\substack{(27.9)\\(27.12)}}{\subset} G_{k_{n_0}}.$$

Таким образом, $M_{k_{n_0}} \in Q \subset s$ и $M_{k_{n_0}} \in G_{k_{n_0}}$, что противоречит выбору точек M_k, $k = 1, 2, \ldots$. □

З а м е ч а н и е. В дальнейшем (в п. 35.3) дано обобщение леммы Гейне—Бореля (см. лемму 4 в п. 23.8*) на плоские и пространственные множества. С помощью этого обобщения доказанная выше лемма может быть доказана значительно проще.

Д о к а з а т е л ь с т в о т е о р е м ы 2. Предварительно заметим, что из условия $G_1 \subset G_2 \subset \ldots G_k \subset \ldots$ следует (см. теорему 1), что

$$\text{mes}\, G_1 \leqslant \text{mes}\, G_2 \leqslant \ldots \leqslant \text{mes}\, G_k, \ldots, \tag{27.13}$$

поэтому последовательность G_k, $k = 1, 2, \ldots$, всегда имеет предел, конечный или равный $+\infty$.

Рассмотрим два случая.

1. Пусть все множества $s_m(G)$, $m = 0, 1, \ldots$, состоят из конечного числа квадратов. В этом случае, по лемме 1, для всякого номера m существует такой номер k_m, что

$$s_m(G) \subset G_{k_m}, \quad m = 1, 2, \ldots . \qquad (27.14)$$

При этом выберем k_m так, что $k_{m'} > k_m$ при $m' > m$. Это всегда можно сделать, например, следующим образом. Если выбраны номера $k_1 < k_2 < \ldots < k_{m-1}$ и для множества $s_m(G)$, согласно лемме 1, найдено множество G_n такое, что

$$s_m(G) \subset G_n, \qquad (27.15)$$

то обозначим через k_m какое-либо натуральное число такое, что $k_m > k_{m-1}$ и $k_m \geqslant n$; тогда $G_n \subset G_{k_m}$ и, значит, $s_m(G) \subset \subset G_{k_m}$. Таким образом, построенная последовательность k_m, $m = 1, 2, \ldots$, является подпоследовательностью последовательности натуральных чисел.

Обозначим теперь через $\tilde{s}_m(G)$ совокупность всех внутренних точек множества $s_m(G)$. Очевидно, $\tilde{s}_m(G)$ — открытое множество и $\tilde{s}_m(G) \subset s_m(G) \subset G_{k_m}$, поэтому, в силу теоремы 1,

$$\operatorname{mes} \tilde{s}_m(G) \leqslant \operatorname{mes} G_{k_m}. \qquad (27.16)$$

Поскольку $G_k \subset G$, $k = 1, 2, \ldots$, то в силу той же теоремы 1,

$$\operatorname{mes} G_{k_m} \leqslant \operatorname{mes} G. \qquad (27.17)$$

Объединяя неравенства (27.16) и (27.17), получаем

$$\operatorname{mes} s_m(G) = \operatorname{mes} \tilde{s}_m(G) \leqslant \operatorname{mes} G_{k_m} \leqslant \operatorname{mes} G.$$

Переходя в этом неравенстве к пределу при $m \to \infty$, имеем

$$\lim_{m \to \infty} \operatorname{mes} G_{k_m} = \operatorname{mes} G,$$

так как, согласно (27.3),

$$\lim_{m \to \infty} \operatorname{mes} s_m(G) = \operatorname{mes} G.$$

Последовательность $\{\operatorname{mes} G_k\}$, как отмечалось выше, имеет конечный или бесконечный предел, поэтому он совпадает с пределом любой ее подпоследовательности, следовательно, $\lim\limits_{k \to \infty} \operatorname{mes} G_k = \operatorname{mes} G$, т. е. выполняется равенство (27.8).

2. Пусть существует множество $s_m(G)$, содержащее бесконечно много квадратов; тогда пл. $s_m(G) = +\infty$, поэтому и $\text{mes}\, G = +\infty$. Покажем, что в этом случае и

$$\lim_{k \to \infty} \text{mes}\, G_k = +\infty. \qquad (27.18)$$

Пусть задано $\varepsilon > 0$ и пусть $s_m(G)$ состоит из бесконечного множества квадратов. Площадь каждого квадрата ранга m равна $\frac{1}{10^{2m}}$. Зафиксируем натуральное число n так, чтобы

$$\frac{n}{10^{2m}} > \varepsilon, \qquad (27.19)$$

и выберем из $s_m(G)$ n каких-либо квадратов. Обозначим множество их точек через s. Тогда

$$\text{пл.}\, s = \frac{n}{10^{2m}}. \qquad (27.20)$$

В силу леммы 1, существует такой номер k, что

$$s \subset G_k. \qquad (27.21)$$

Обозначим через $\tilde{s}$ множество внутренних точек многоугольника s. Согласно теореме 1 и формулам (27.19), (27.20), получим: $\text{mes}\, G_k \geqslant \text{пл.}\, \tilde{s} = \text{пл.}\, s > \varepsilon$. В силу же (27.13), и для всех $k' \geqslant k$ имеем $\text{mes}\, G_{k'} > \varepsilon$. Это и означает, что $\lim\limits_{k \to \infty} \text{mes}\, G_k = +\infty$, т. е. выполнение условия (27.18). □

Примером неограниченной плоской области, имеющей бесконечную меру, является полоса $G = \{(x, y)\colon 0 < y < 1\}$. Она содержит в себе бесконечное множество, например, квадратов первого ранга, поэтому $\text{mes}\, G = +\infty$.

В качестве примера неограниченной области с конечной площадью рассмотрим фигуру, построенную Н. Оресмом[1]. Пусть Q — единичный квадрат:

$$Q = \{(x, y)\colon 0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant 1\}.$$

Положим

$$G_1 = \left\{(x, y)\colon 0 < x < 1,\ 0 < y < \frac{1}{2}\right\},$$

$$G_2 = G_1 \cup \left\{(x, y)\colon 1 \leqslant x < 2,\ 0 < y < \frac{1}{4}\right\},$$

[1] Н. Оресм (ок. 1323—1382) — французский математик.

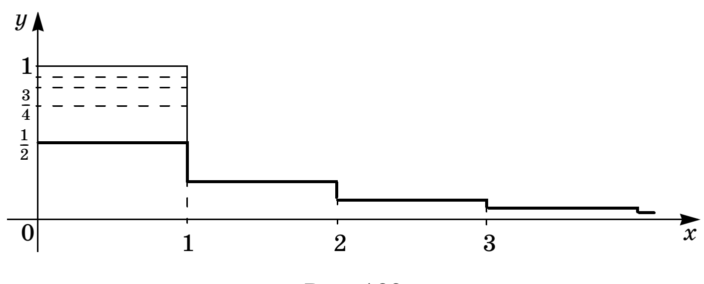


Рис. 123

вообще,

$$G_{k+1} = G_k \cup \left\{(x, y)\colon k \leqslant x < k + 1,\ 0 < y < \frac{1}{2^{k+1}}\right\},\ k = 1, 2, \ldots .$$

Каждое множество G_k открыто (почему?).

Наглядно образование множеств G_k можно представить следующим образом: G_1 — половина квадрата Q; для получения G_2 берется половина оставшейся половины квадрата Q и прикладывается соответствующим образом к G_1, получается G_2; далее, половина оставшейся части квадрата Q прикладывается уже к Q_3 (рис. 123) и т. д. Очевидно, имеем цепочку включений $G_1 \subset G_2 \subset \ldots \subset G_k \subset \ldots$ и

$$\text{пл. } G_k = \frac{1}{2} + \frac{1}{2^2} + \ldots + \frac{1}{2^k} = \frac{\frac{1}{2} - \frac{1}{2^{k+1}}}{1 - \frac{1}{2}} = 1 - \frac{1}{2^k}.$$

Положим $G = \bigcup\limits_{k=1}^{\infty} G_k$. Множество G открыто и не ограничено. Найдем, применив теорему 2, его площадь:

$$\operatorname{mes} G = \lim_{k \to \infty} \operatorname{mes} G_k = \lim_{k \to \infty} \left(1 - \frac{1}{2^k}\right) = 1.$$

Мера mes (объем) открытых множеств в пространстве определяется с помощью аналогичной конструкции, следует только исходить не из разбиений плоскости на квадраты (*квадрильяжей*), а из разбиений пространства на соответствующие кубы (*кубильяжи*). На пространственный случай переносятся и теоремы, доказанные в этом параграфе. Мы еще вернемся к рассмотрению меры множеств в дальнейших главах (см. п. 44.1). Там будут изложены более полно свойства

меры (например, ее поведение при объединении множеств — аддитивность меры).

УПРАЖНЕНИЯ. 1. Доказать, что площадь прямоугольника равна произведению его сторон.

2. Пусть G — прямой круговой цилиндр, основанием которого является круг K, а высота имеет длину h. Доказать, что $\operatorname{mes} G = h \operatorname{mes} K$, где $\operatorname{mes} G$ есть мера (объем) G в пространстве, а $\operatorname{mes} K$ мера (площадь) K на плоскости.

§ 28.

Некоторые геометрические и физические приложения определенного интеграла

28.1. Вычисление площадей

В этом пункте будут выведены формулы для вычисления площадей некоторых плоских областей. При этом воспользуемся известными из элементарной математики свойствами площади простейших плоских фигур (многоугольников, секторов), например тем, что при объединении таких фигур, не имеющих общих внутренних точек, их площади складываются. Отметим, что это утверждение будет доказано в п. 44.1.

ТЕОРЕМА 1. *Пусть функция f определена, неотрицательна и непрерывна на отрезке $[a, b]$. Тогда площадь S множества*

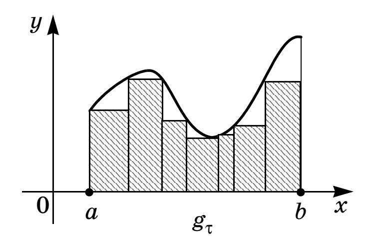
$$G = \{(x, y)\colon a < x < b,\ 0 < y < f(x)\}$$

выражается формулой

$$S = \int_a^b f(x)\, dx. \tag{28.1}$$

Множество G является открытым ограниченным множеством. Действительно, его ограниченность следует из того, что функция f, будучи непрерывной на отрезке $[a, b]$, ограничена на нем.

Покажем, что множество G открыто. Пусть $(x_0, y_0) \in G$; тогда $a < x_0 < b$ и $0 < y_0 < f(x_0)$. Возьмем какое-либо число $\eta > 0$ такое, что $0 < y_0 - \eta < y_0 < y_0 + \eta < f(x_0)$. В силу непре-



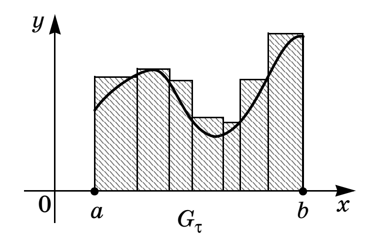


Рис. 124

рывности функции f в точке x_0 существует такое $\delta > 0$, что и для всех $x \in (x_0 - \delta, x_0 + \delta)$ выполняется неравенство $y_0 - \eta < f(x) < y_0 + \eta$. Поэтому, если $\varepsilon = \min\{\delta; \eta\}$, то ε — окрестность $U((x_0, y_0); \varepsilon)$ точки (x_0, y_0) содержится в множестве G, т. е. (x_0, y_0) — его внутренняя точка. Это и означает, что G — открытое множество.

Множество G обычно называют *криволинейной трапецией*, порожденной графиком функции f (см. рис. 106).

Множество $\overline{G} \stackrel{\text{def}}{=} \{(x, y)\colon a \leqslant x \leqslant b,\ 0 \leqslant y \leqslant f(x)\}$ называется замкнутой криволинейной трапецией. Площадь криволинейной трапеции G называется и площадью замкнутой криволинейной трапеции $\overline{G}$ (естественность этого определения показана в п. 44.1).

Доказательство. Пусть $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ — некоторое разбиение отрезка $[a, b]$. Обозначим через G_τ и g_τ замкнутые многоугольники, составленные из всех прямоугольников вида

$$G_{\tau, i} = \{(x, y\} : x_{i-1} \leqslant x \leqslant x_i,\ 0 \leqslant y \leqslant M_i\},$$
$$g_{\tau, i} = \{(x, y) : x_{i-1} \leqslant x \leqslant x_i,\ 0 \leqslant y \leqslant m_i\},$$

где $m_i = \inf\limits_{x_{i-1} \leqslant x \leqslant x} f(x)$, $M_i = \sup\limits_{x_{i-1} \leqslant x \leqslant x_i} f(x)$, $i = 1, 2, \ldots, i_\tau$.

Таким образом (рис. 124),

$$G_\tau = \bigcup_{i=1}^{i_\tau} G_{\tau, i}, \quad g_\tau = \bigcup_{i=1}^{i_\tau} g_{\tau, i}. \qquad (28.2)$$

Если обозначить через $\tilde{G}_\tau$ и $\tilde{g}_\tau$ множество внутренних точек многоугольников G_τ и g_τ, то

$$\tilde{g}_\tau \subset G \subset \tilde{G}_\tau. \qquad (28.3)$$

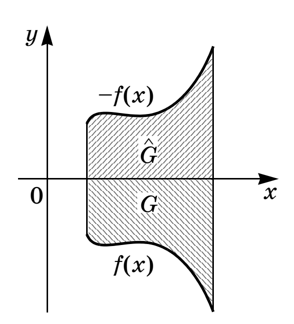


Рис. 125

Если S_τ и s_τ — соответственно верхняя и нижняя суммы Дарбу функции f на отрезке $[a, b]$, соответствующие его разбиению τ, то очевидно, что пл. $\tilde{g}_\tau = s_\tau$, пл. $\tilde{G}_\tau = S_\tau$. Поэтому из (28.3), в силу монотонности меры, следует, что

$$s_\tau \leqslant \text{mes}\, G \leqslant S_\tau. \tag{28.4}$$

Так как

$$\lim_{|\tau| \to 0} s_\tau = \lim_{|\tau| \to 0} S_\tau = \int_a^b f(x)\, dx, \tag{28.5}$$

то отсюда в силу неравенства (28.4) следует, что

$$\text{mes}\, G = \int_a^b f(x)\, dx. \ \square$$

Как известно (см. п. 23.5),

$$\lim_{|\tau| \to 0} \sigma_\tau = \lim_{|\tau| \to 0} s_\tau = \lim_{|\tau| \to 0} S_\tau = \int_a^b f(x)\, dx,$$

поэтому, в силу формулы (28.1),

$$\lim_{|\tau| \to 0} \sigma_\tau = \lim_{|\tau| \to 0} s_\tau = \lim_{|\tau| \to 0} S_\tau = \text{mes}\, G.$$

Таким образом, геометрически интегральные суммы Римана и суммы Дарбу равны приближенному значению площади рассматриваемой криволинейной трапеции, причем любая точность достигается выбором достаточной мелкости разбиения τ, а предел интегральных сумм равен истинному значению указанной площади.

Пусть теперь функция f непрерывна и неположительна на отрезке $[a, b]$. Положим в этом случае

$$G = \{(x, y) : a < x < b,\ f(x) < y < 0\}.$$

Пусть $\hat{G}$ — множество, симметричное множеству G относительно оси Ox[1] (рис. 125); тогда

$$\text{mes}\, \hat{G} = \text{mes}\, G. \tag{28.6}$$

В рассматриваемом случае функция $-f$ неотрицательна на отрезке $[a, b]$, поэтому

$$\text{mes}\, \hat{G} = \int_a^b [-f(x)]\, dx = -\int_a^b f(x)\, dx. \tag{28.7}$$

[1] Это означает, что $\hat{G} = \{(x, y) : (x, -y) \in G\}$.

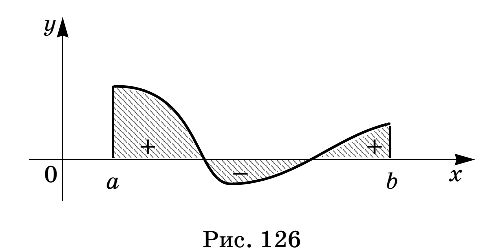


Рис. 126

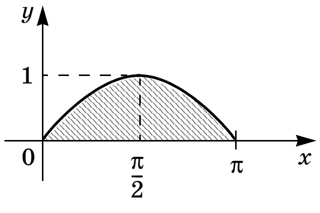


Рис. 127

Сравнив (28.6) и (28.7), получим

$$\operatorname{mes} G = -\int_a^b f(x)\,dx,$$

т. е. здесь интеграл $\int_a^b f(x)\,dx$ равен, с точностью до знака, значению площади криволинейной трапеции.

Если же функция f меняет знак на отрезке $[a, b]$ в конечном числе точек, то интеграл $\int_a^b f(x)\,dx$ равен алгебраической сумме площадей соответствующих криволинейных трапеций, ограниченных частями графика функции f, отрезками оси Ox и, быть может, отрезками, параллельными оси Oy (рис. 126).

Как видно, одной из задач, естественным образом приводящих к понятию определенного интеграла, является задача вычисления площадей. Развитый аппарат интегрального исчисления дает общий и единый метод вычисления площадей разнообразных плоских фигур.

Примеры. 1. Найдем площадь S, ограниченную отрезком оси Ox и одной аркой синусоиды (рис. 127):

$$S = \int_0^\pi \sin x\,dx = -\cos x\Big|_0^\pi = 2.$$

Здесь, как и всегда в дальнейшем, говоря о множестве, ограниченном некоторой кривой, являющейся простым замкнутым контуром (см. п. 16.1), называемым границей этого множества, всегда будем иметь в виду ограниченное множество. Всякое неограниченное множество, границей которого служит подобный контур, будем называть внешним (для данного контура). В рассматриваемом случае внешним множеством является «внешность» множества, заштрихованного на рисунке 127, т. е. совокупность точек плоскости, не принадлежащих заштрихованному множеству. Внешнее множество

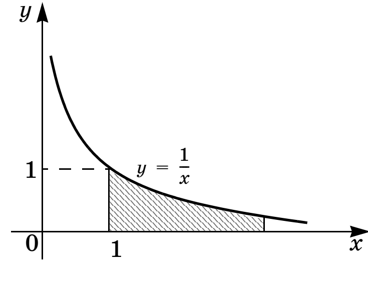


Рис. 128

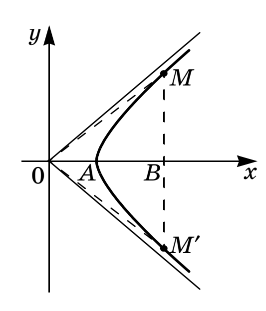


Рис. 129

всегда имеет бесконечную площадь. Действительно, всякая кривая ограничена (см. п. 16.3), поэтому во внешнем множестве любого простого контура содержится, например, квадрат со сколь угодно большой стороной. Отсюда сразу и следует бесконечность площади внешнего множества.

2. Найдем площадь S множества, ограниченного гиперболой $y = \frac{1}{x}$ отрезком оси Ox, отрезком прямой $x = 1$ и отрезком прямой, проходящей через точку оси Ox с абсциссой, равной x и параллельной оси ординат (рис. 128):

$$S = \int_1^x \frac{dt}{t} = \ln\Big|_1^x = \ln x.$$

3. Рассмотрим на правой ветви гиперболы $x^2 - y^2 = 1$ (рис. 129) точки $M = (x, y)$ и $M' = (x, -y)$ $(x > 0)$ и выясним, как зависит площадь s сектора OMM' этой гиперболы от абсциссы x точек M и M'. Прежде всего заметим, что функция $s = s(x)$, принимая значения, равные площадям некоторых фигур, неотрицательна: $s(x) \geqslant 0$ при $x > 0$.

Пусть $BM = y$. Имеем (см. рис. 129)

$$s = s(x) = \text{пл. } OMM' = 2\text{пл. } OMA = 2(\text{пл. } \Delta OMB - \text{пл. } AMB) =$$
$$= 2\left(\frac{1}{2} OB \cdot MB - \int_1^x \sqrt{t^2 - 1}\, dt\right) = x\sqrt{x^2 - 1} - 2\int_1^x \sqrt{t^2 - 1}\, dt.$$

Для вычисления получившегося интеграла сделаем замену переменного $t = \operatorname{ch} u$:

$$2\int_0^x \sqrt{t^2 - 1}\, dt = 2\int_0^{\operatorname{Areach} x} \operatorname{sh}^2 u\, du = \int_0^{\operatorname{Areach} x} (\operatorname{ch} 2u - 1)\, du =$$
$$= \frac{1}{2}\operatorname{sh} 2\operatorname{Areach} x - \operatorname{Areach} x =$$
$$= \operatorname{sh} \operatorname{Areach} x \cdot \operatorname{ch} \operatorname{Areach} x - \operatorname{Areach} x = x\sqrt{x^2 - 1} - \operatorname{Areach} x.$$

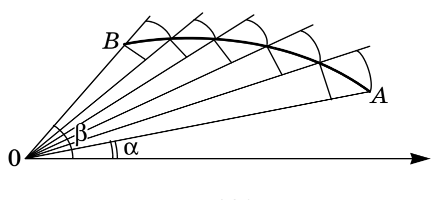


Рис. 130

Рис. 131

Поэтому

$$s = s(x) = \operatorname{Areach} x,$$

откуда $x = \operatorname{ch} s$. Таким образом, площадь сектора гиперболы равна функции, обратной гиперболическому косинусу $\operatorname{ch} s$ при $s \geqslant 0$.

В силу этого, при параметрическом представлении $x = \operatorname{ch} s$, $y = \operatorname{sh} s$ гиперболы $x^2 - y^2 = 1$ параметр $s = \operatorname{Areach} x = \operatorname{Areach} y$ совпадает со значением площади соответствующего сектора гиперболы, взятой с надлежащим знаком (площадям секторов левой ветви гиперболы $x^2 - y^2 = 1$ приписывается знак минус).

4. Вычислим площадь, ограниченную эллипсом $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Так как лежащий выше оси абсцисс полуэллипс описывается уравнением $y = \frac{b}{a}\sqrt{a^2 - x^2}$, то для четверти искомой площади S имеем (см. пример 5 в п. 18.4 или пример 1 в п. 18.5)

$$\frac{1}{4}S = \frac{b}{a}\int_0^a \sqrt{a^2 - x^2}\,dx = \frac{b}{a}\left(\frac{a^2}{2}\arcsin\frac{x}{a} + \frac{x}{2}\sqrt{a^2 - x^2}\right)\Big|_0^a = \frac{ab\pi}{4},$$

откуда $S = \pi ab$.

Формула для площади в полярных координатах. Найдем теперь формулу для площади сектора кривой, заданной уравнением, связывающим ее полярные координаты: $\rho = \rho(\varphi)$, где $\rho = \rho(\varphi)$ — неотрицательная, непрерывная на отрезке $[\alpha, \beta]$ функция, $0 \leqslant \alpha \leqslant \beta \leqslant 2\pi$. Пусть G — открытое множество, ограниченное кривой $\overset{\frown}{AB}$, описываемой в полярных координатах уравнением $\rho = \rho(\varphi)$ и, быть может, отрезками OA и OB лучей $\varphi = \alpha$ и $\varphi = \beta$ (рис. 130), $G = \{(\rho, \varphi) : \alpha < \varphi < \beta, 0 < \rho < \rho(\varphi)\}$.

Пусть $\tau = \{\varphi_i\}_{i=0}^{i=k}$ — некоторое разбиение отрезка $[\alpha, \beta]$. Положим

$$\Delta\varphi_i = \varphi_i - \varphi_{i-1},\ m_i = \inf_{\varphi_{i-1} \leqslant \varphi \leqslant \varphi_i} \rho(\varphi),\ M_i = \sup_{\varphi_{i-1} \leqslant \varphi \leqslant \varphi_i} \rho(\varphi),$$

$$g_{i,\tau} = \{(\rho, \varphi) : \varphi_{i-1} \leqslant \varphi \leqslant \varphi_i,\ 0 \leqslant \rho \leqslant m_i\},$$

$$G_{i,\tau} = \{(\rho, \varphi) : \varphi_{i-1} \leqslant \varphi \leqslant \varphi_i,\ 0 \leqslant \rho \leqslant M_i\}.$$

Впишем во множество G и опишем вокруг него ступенчатые фигуры g_τ и G_τ, составленные из круговых секторов $g_{i,\tau}$ и $G_{i,\tau}$, $i = 1, 2, \ldots, k$:

$$g_\tau = \bigcup_{i=1}^{k} g_{i,\tau},\quad G_\tau = \bigcup_{i=1}^{k} G_{i,\tau}.$$

Обозначим через $\tilde{g}_\tau$ и $\tilde{G}_\tau$ совокупности всех внутренних точек множеств g_τ и G_τ. Очевидно, $\tilde{g}_\tau$ и $\tilde{G}_\tau$ — открытые множества и $\tilde{g}_\tau \subset G \subset \tilde{G}_\tau$, поэтому, согласно свойству монотонности площади,

$$\text{пл. } \tilde{g}_\tau \leqslant \operatorname{mes} G \leqslant \text{пл. } \tilde{G}_\tau.$$

Но пл. $\tilde{g}_\tau$ = пл. g_τ, пл. $\tilde{G}_\tau$ = пл. G_τ, следовательно,

$$\text{пл. } g_\tau \leqslant \operatorname{mes} G \leqslant \text{пл. } G_\tau. \tag{28.8}$$

Площади круговых секторов $g_{i,\tau}$ и $G_{i,\tau}$ равны соответственно $\frac{1}{2} m_i^2 \Delta\varphi_i$ и $\frac{1}{2} M_i^2 \Delta\varphi_i$ (рис. 131). Из элементарной математики известно, что при объединении плоских фигур их площади складываются (см. об этом также в п. 44.1), поэтому

$$\text{пл. } g_\tau = \frac{1}{2} \sum_{i=1}^{i_\tau} m_i^2 \Delta\varphi_i,\quad \text{пл. } G_\tau = \frac{1}{2} \sum_{i=1}^{i_\tau} M_i^2 \Delta\varphi_i.$$

Из этих равенств видно, что пл. g_τ и пл. G_τ являются соответственно нижней и верхней суммами Дарбу для функции $\frac{1}{2}\rho^2(\varphi)$ на отрезке $[\alpha, \beta]$: s_τ = пл. g_τ, S_τ = пл. G_τ, следовательно,

$$s_\tau \leqslant \frac{1}{2} \int_\alpha^\beta \rho^2(\varphi)\, d\varphi \leqslant S_\tau.$$

Вычитая это неравенство из неравенства (28.8), переписанного в виде $S_\tau \geqslant \operatorname{mes} G \geqslant s_\tau$, получаем

$$s_\tau - S_\tau \leqslant \operatorname{mes} G - \frac{1}{2} \int_\alpha^\beta \rho^2(\varphi)\, d\varphi \leqslant S_\tau - s_\tau.$$

Отсюда, перейдя к пределу при $|\tau| \to 0$, имеем

$$\text{mes}\, G = \frac{1}{2}\int_{\alpha}^{\beta} \rho^2(\varphi)\, d\varphi. \ \square \qquad (28.9)$$

5. Найдем площадь S фигуры, ограниченной кардиоидой $\rho = a(1 + \cos\varphi)$ (см. п. 17.5), которая изображена на рисунке 132. По формуле (28.9) получим

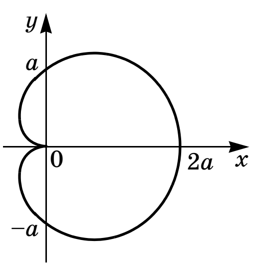


Рис. 132

$$S = \frac{a^2}{2}\int_0^{2\pi} (1 + \cos\varphi)^2\, d\varphi =$$

$$= \frac{a^2}{2}\int_0^{2\pi} d\varphi + a^2\int_0^{2\pi} \cos\varphi\, d\varphi + \frac{a^2}{2}\int_0^{2\pi} \frac{1 + \cos 2\varphi}{2}\, d\varphi = \frac{3}{2}\pi a^2.$$

28.2*. Интегральные неравенства Гёльдера и Минковского

Установим теперь справедливость двух интегральных неравенств, нередко применяющихся в математическом анализе. Предварительно докажем следующие три леммы.

ЛЕММА 1. *Если функция f ограничена на отрезке* $[a, b]$:

$$|f(x)| \leqslant c, \ \ a \leqslant x \leqslant b, \qquad (28.10)$$

число $p \geqslant 1$, а $\omega(|f|)$ и $\omega(|f|^p)$ являются колебаниями на отрезке $[a, b]$ соответственно функций $|f|$ и $|f|^p$, то имеет место неравенство

$$\omega(|f|^p) \leqslant pc^{p-1}\omega(|f|). \qquad (28.11)$$

Доказательство. Пусть $\alpha \geqslant 0$, $\beta \geqslant 0$, $\gamma = \max\{\alpha; \beta\}$, $p \geqslant 1$. Покажем, что имеет место числовое неравенство

$$|\beta^p - \alpha^p| \leqslant p\gamma^{p-1}|\beta - \alpha|. \qquad (28.12)$$

Рассмотрим функцию $\varphi(t) = t^p$. В силу примененной к ней на отрезке с концами в точках α и β формулы конечных приращений Лагранжа, имеем

$$|\beta^p - \alpha^p| = |\varphi(\beta) - \varphi(\alpha)| =$$

$$= |\varphi'(\xi)|\,|\beta - \alpha| = p\xi^{p-1}|\beta - \alpha|, \qquad (28.13)$$

где точка ξ лежит между точками α и β, поэтому

$$0 \leqslant \xi \leqslant \max\{\alpha; \beta\} = \gamma. \qquad (28.14)$$

Из формул (28.13) и (28.14) сразу следует неравенство (28.12).

Если f — функция, определенная на отрезке $[a, b]$, $x \in [a, b]$, $x' \in [a, b]$, то, положив в неравенстве (28.12) $\alpha = |f(x)|$, $\beta = |f(x')|$ и заметив, что $\max\{|f(x)|, |f(x')|\} \underset{(28.10)}{\leqslant} c$, получим

$$\left||f(x')|^p - |f(x)|^p\right| \leqslant pc^{p-1}\left||f(x')| - |f(x)|\right|.$$

Откуда

$$\omega(|f|^p) = \sup_{x, x' \in [a, b]} \left||f(x')|^p - |f(x)|^p\right| \leqslant$$

$$\leqslant pc^{p-1} \sup_{x, x' \in [a, b]} \left||f(x')| - |f(x)|\right| = pc^{p-1}\omega(|f|).$$

Неравенство (28.11) доказано. □

ЛЕММА 2. *Если функция f интегрируема на некотором отрезке и число $p \geqslant 1$, то функция $|f|^p$ также интегрируема на этом отрезке.*

Доказательство. Пусть функция f интегрируема на отрезке $[a, b]$, тогда она ограничена (теорема 2 в п. 23.3):

$$c = \sup_{x \in [a, b]} |f(x)| < +\infty.$$

Если $\tau = \{x_i\}_{i=0}^{i=i_\tau}$ — какое-либо разбиение отрезка $[a, b]$, то на любом отрезке $[x_{i-1}, x_i]$ колебания $\omega_i(|f|)$ и $\omega_i(|f|^p)$ соответственно функций $|f|$ и $|f|^p$ в силу неравенства (28.11) удовлетворяет неравенству

$$\omega_i(|f|^p) \leqslant pc^{p-1}\omega_i(|f|).$$

Поэтому

$$0 \leqslant \sum_{i=1}^{i_\tau} \omega_i(|f|^p)\Delta x_i \leqslant pc^{p-1} \sum_{i=1}^{i_\tau} \omega_i(|f|)\Delta x_i, \qquad (28.15)$$

$$\Delta x_i = x_i - x_{i-1},\ i = 1, 2, \ldots, i_\tau.$$

Из интегрируемости функции f, согласно свойству 10^0 интеграла (см. п. 24.1), следует, что ее абсолютная величина также интегрируема, следовательно, из критерия интегрируемости функции, сформулированного в следствии 2 теоремы 3 в п. 23.5, явствует, что

$$\lim_{|\tau| \to 0} \sum_{i=1}^{i_\tau} \omega_i(|f|)\Delta x_i = 0.$$

Поэтому в силу неравенства (28.15) имеем

$$\lim_{|\tau| \to 0} \sum_{i=1}^{i_\tau} \omega_i(|f|^p)\Delta x_i = 0.$$

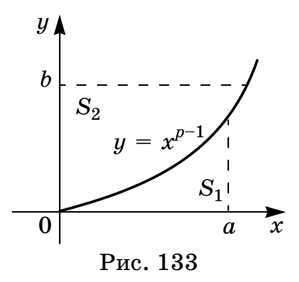


Рис. 133

А это, согласно тому же критерию интегрируемости функции, означает, что функция $|f|^p$ интегрируема на отрезке $[a, b]$. □

Установим теперь, используя геометрический смысл интеграла (см. п. 28.1), одно необходимое для дальнейшего числовое неравенство.

ЛЕММА 3. *Пусть p и q — такие действительные числа, что*

$$\frac{1}{p} + \frac{1}{q} = 1,\ \ p > 1, \tag{28.16}$$

тогда для произвольных $a \geqslant 0$ и $b \geqslant 0$ имеет место неравенство

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}. \tag{28.17}$$

Доказательство. Рассмотрим кривую $y = x^{p-1}$ или, что то же самое, кривую $x = y^{q-1}$, так как

$$(p - 1)(q - 1) = 1.$$

Выберем произвольно $a \geqslant 0$ и $b \geqslant 0$ и подсчитаем площади S_1 и S_2 (рис. 133):

$$S_1 = \int_0^a x^{p-1}\,dx = \frac{a^p}{p},\ \ S_2 = \int_0^b y^{q-1}\,dy = \frac{b^q}{q}. \tag{28.18}$$

Поскольку прямоугольник со сторонами a и b содержится в объединении рассматриваемых замкнутых криволинейных трапеций (см. п. 28.1), площади которых равны S_1 и S_2, площадь этого прямоугольника не превышает суммы $S_1 + S_2$, т. е. $ab \leqslant S_1 + S_2$. Подставив в это неравенство выражения (28.18) для S_1 и S_2, получим неравенство (28.17). При этом ясно, что $ab = S_1 + S_2$ в том и только том случае, когда $b = a^{p-1}$. □

Перейдем к доказательству интегральных неравенств, указанных в заглавии. Пусть функции f и g определены и интегрируемы на отрезке $[a, b]$, $1 < p < +\infty$, а число q определяется равенством

$$\frac{1}{p} + \frac{1}{q} = 1 \tag{28.19}$$

(см. (28.16)). Тогда имеем

$$\int_a^b |f(x)g(x)|\,dx \leqslant \left[\int_a^b |f(x)|^p\,dx\right]^{\frac{1}{p}} \left[\int_a^b |g(x)|^q\,dx\right]^{\frac{1}{p}} \qquad (28.20)$$

(*неравенство Гёльдера*[1]),

$$\left[\int_a^b |f(x)+g(x)|^p\right]^{\frac{1}{p}} \leqslant \left[\int_a^b |f(x)|^p\,dx\right]^{\frac{1}{p}} + \left[\int_a^b |g(x))|^p\,dx\right]^{\frac{1}{p}} \qquad (28.21)$$

(*неравенство Минковского*[2]). Докажем эти неравенства. Введем для краткости обозначения

$$\|f\|_p \stackrel{\text{def}}{=} \left[\int_a^b |f(x)|^p\,dx\right]^{\frac{1}{p}},\ \ \|g\|_q \stackrel{\text{def}}{=} \left[\int_a^b |g(x)|^q\,dx\right]^{\frac{1}{q}}. \qquad (28.22)$$

В неравенстве (28.17)

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q},\ \ a \geqslant 0,\, b \geqslant 0,$$

положим

$$a = \frac{|f(x)|}{\|f\|_p},\ \ b = \frac{|g(x)|}{\|g\|_q},\ x \in [a, b].$$

Тогда для любого $x \in [a, b]$

$$\frac{|f(x)|}{\|f\|_p}\frac{|g(x)|}{\|g\|_q} \leqslant \frac{1}{p}\frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{q}\frac{|q(x)|^q}{\|q\|_q^q}.$$

Проинтегрировав это неравенство по отрезку $[a, b]$ и использовав формулы (28.19) и (28.22), найдем

$$\frac{1}{\|f\|_p\|q\|_q}\int_a^b |f(x)g(x)|\,dx \leqslant \frac{1}{p\|f\|_p^p}\int_a^b |f(x)|^p\,dx + \frac{1}{q\|g\|_q^q}\int_a^b |g(x)|^q\,dx =$$
$$= \frac{1}{p} + \frac{1}{q} = 1.$$

Поэтому

$$\int_a^b |f(x)g(x)|\,dx \leqslant \|f\|_p\,\|g\|_q,$$

т. е. неравенство (28.20) доказано.

Докажем неравенство (28.21). Легко убедиться в справедливости неравенства

$$\int_a^b |f(x)+g(x)|^p\,dx = \int_a^b |f(x)+g(x)|\,|f(x)+g(x)|^{p-1}\,dx \leqslant$$
$$\leqslant \int_a^b |f(x)|\,|f(x)+g(x)|^{p-1}\,dx + \int_a^b |g(x)|\,|f(x)+g(x)|^{p-1}\,dx.$$

---

[1] О. Л. Гёльдер (1859—1937) — немецкий математик.

[2] Г. Минковский (1864—1906) — математик, родился в России, работал в Швейцарии и Германии.

Применив к каждому из полученных интегралов неравенство Гёльдера и заметив, что $q(p-1)=p$ (см. 28.19), получим:

$$\int_a^b |f(x)+g(x)|^p\,dx \leqslant \left[\int_a^b |f(x)|^p\,dx\right]^{1/p}\left[\int_a^b |f(x)+g(x)|^{q(p-1)}\,dx\right]^{1/q} +$$

$$+\left[\int_a^b |g(x)|^p\,dx\right]^{1/p}\left[\int_a^b |f(x)+g(x)|^{q(p-1)}\,dx\right]^{1/q} =$$

$$=\left\{\left[\int_a^b |f(x)|^p\,dx\right]^{1/p}+\left[\int_a^b |g(x)|^p\,dx\right]^{1/p}\right\}\left[\int_a^b |f(x)+g(x)|^p\,dx\right]^{1/q}. \tag{28.23}$$

Если левая часть этого неравенства равна нулю, то неравенство (28.21), очевидно, справедливо, если же она не равна нулю, то, сократив обе части неравенства (28.23) на множитель $\left[\int_a^b |f(x)+g(x)|^p\,dx\right]^{1/q}$, в силу соотношения (28.19) получим неравенство Минковского. □

Отметим важный частный случай неравенства Гёльдера. При $p=q=2$ имеем

$$\int_a^b |f(x)g(x)|\,dx \leqslant \sqrt{\int_a^b |f(x)|^2dx}\,\sqrt{\int_a^b |g(x)|^2dx} \tag{28.24}$$

(*неравенство Коши—Буняковского*[1]).

З а м е ч а н и е. Отметим, что при доказательстве неравенства Минковского (28.21) была использована интегрируемость функции $|f(x)+g(x)|^{p-1}$, $p>1$. Здесь может оказаться, что $0<p-1<1$ и поэтому этот факт требует доказательства. Интегрируемость функции $|f(x)+g(x)|^p$, $p>0$, следует, например, из критерия интегрируемости Лебега (см. п. 23.10*), так как множества точек разрыва у функций $|f(x)+g(x)|$ и $|f(x)+g(x)|^p$ совпадают при любом $p>0$.

Можно идти и другим путем, показав для колебаний функций справедливость неравенства

$$\omega(|f|^p;[a,b]) \leqslant \omega^p(|f|;[a,b]),\ 0<p<1.$$

---

[1] В. Я. Буняковский (1804—1899) — русский математик.

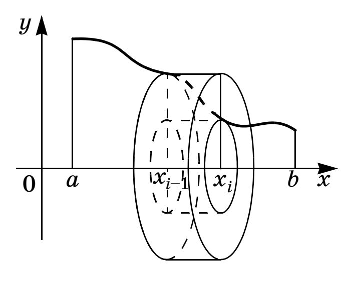


Рис. 134

Однако доказать с помощью этого неравенства выполнение условия

$$\lim_{|\tau|\to 0}\sum_{i=1}^{i_\tau}\omega_i(|f|^p)\Delta x_i = 0$$

(см. следствие 2 теоремы 3 в п. 23.5) затруднительно по причине недостаточности имеющихся пока знаний: здесь целесообразно использовать неравенство Минковского для сумм, которое будет доказано в дальнейшем (в п. 30.8*).

28.3. Объем тела вращения

В конце п. 27.2 отмечалось, что понятие объема в пространстве вводится аналогично понятию площади на плоскости. Выведем формулу для вычисления объемов тел вращения.

ТЕОРЕМА 2. *Пусть функция $f(x)$ неотрицательна и непрерывна на отрезке $[a, b]$, а Q — тело, полученное вращением криволинейной трапеции G, порожденной графиком функции f вокруг оси Ox. Тогда для его объема* $\operatorname{mes} Q$ *справедлива формула*

$$\operatorname{mes} Q = \pi\int_a^b f^2(x)\,dx. \tag{28.25}$$

Доказательство. Обозначим через q_τ и Q_τ тела, образованные вращением вокруг оси Ox ступенчатых фигур $\tilde{g}_\tau$ и $\tilde{G}_\tau$ (см. доказательство теоремы 1). Из включения (28.3) следует, что $q_\tau \subset Q \subset Q_\tau$, поэтому, если $v_\tau = \operatorname{mes} q_\tau$, $V_\tau = \operatorname{mes} Q_\tau$, то

$$v_\tau \leqslant \operatorname{mes} Q \leqslant V_\tau. \tag{28.26}$$

Объемы v_τ и V_τ множеств q_τ и Q_τ равны суммам объемов цилиндров, образованных вращением прямоугольников $g_{\tau, i}$ и $G_{\tau, i}$ (рис. 134):

$$v_\tau = \operatorname{mes} q_\tau = \sum_{i=1}^{i_\tau}\pi m_i^2\Delta x_i, \quad V_\tau = \operatorname{mes} Q_\tau = \sum_{i=1}^{i_\tau}\pi M_i^2\Delta x_i.$$

Из этих равенств видно, что v_τ и V_τ являются нижними и верхними суммами Дарбу функции $\pi f^2(x)$, а так как функция f^2 непрерывна и, следовательно, интегрируема, то

$$v_\tau \leqslant \pi\int_a^b f^2(x)\,dx \leqslant V_\tau \tag{28.27}$$

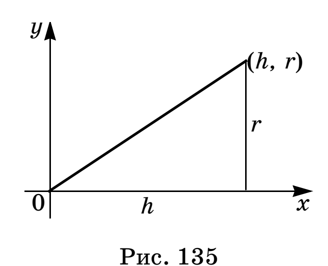


Рис. 135

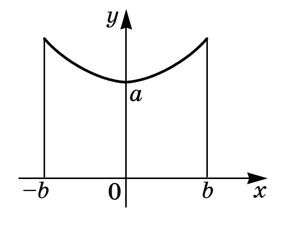


Рис. 136

и

$$\lim_{|\tau| \to 0} [V_\tau - v_\tau] = 0. \tag{28.28}$$

Из неравенств (28.26) и (28.27) следует, что

$$v_\tau - V_\tau \leqslant \pi \int_a^b f^2(x)\,dx - \operatorname{mes} Q \leqslant V_\tau - v_\tau,$$

откуда, в силу (28.28), и вытекает формула (28.25). □

Примеры. 1. Найдем объем V шара радиуса r. Рассматривая этот шар как тело, образованное вращением полуокружности $y = \sqrt{r^2 - x^2}$, $-r \leqslant x \leqslant r$ вокруг оси Ox, по формуле (28.25) получим

$$V = \pi \int_{-r}^{r} (r^2 - x^2) dx = \pi r^2 x \Big|_{-r}^{r} - \frac{\pi x^3}{3}\Big|_{-r}^{r} = 2\pi r^3 - \frac{2}{3}\pi r^3 = \frac{4}{3}\pi r^3.$$

2. Найдем объем V прямого кругового конуса с высотой, равной h, и радиусом основания r. Рассматривая указанный конус как тело, полученное вращением треугольника с вершинами в точках (0, 0), $(h, 0)$ и (h, r) вокруг оси Ox (рис. 135), т. е. вращением «криволинейной трапеции», порожденной графиком функции $y = \frac{r}{h}x$, $0 \leqslant x \leqslant h$, получим, согласно формуле (28.25),

$$V = \frac{\pi r^2}{h^2} \int_0^h x^2\,dx = \frac{\pi r^2 x^3}{3h^2}\Big|_0^h = \frac{\pi r^2 h}{3}.$$

3. Найдем объем V тела, полученного вращением вокруг оси Ox криволинейной трапеции, образованной графиком функции $y = a \operatorname{ch} \frac{x}{a}$, $-b \leqslant x \leqslant b$, называемым *цепной линией* (рис. 136). По формуле (28.25) имеем

$$V = \pi a^2 \int_{-b}^{b} \operatorname{ch}^2 \frac{x}{a}\,dx = \frac{\pi a^2}{2} \int_{-b}^{b} \left(1 + \operatorname{ch} \frac{2x}{a}\right) dx =$$

$$= \left(\frac{\pi a^2 x}{2} + \frac{\pi a^3}{4} \operatorname{sh} \frac{2x}{a}\right)\Big|_{-b}^{b} = \pi a^2 b + \frac{\pi a^3}{2} \operatorname{sh} \frac{2b}{a}.$$

Из рассмотренных в этом параграфе примеров уже отчетливо видны сила и общность методов интегрального исчисления: единым методом быстро и просто получаются формулы для площадей и объемов, как известные ранее из курса элементарной математики, так и совершенно новые. В ближайших пунктах мы рассмотрим еще ряд задач, также легко решаемых методами интегрального исчисления.

28.4. Вычисление длины кривой

Мы рассмотрели ряд задач, приводящих к понятию определенного интеграла. Все они имеют то общее, что в них нахождение значения какой-то величины приводилось к определению предела некоторой интегральной суммы при стремлении мелкости разбиения к нулю, т. е. к определенному интегралу.

Существует, однако, и другой круг задач, приводящих к понятию определенного интеграла. В них известна скорость изменения одной величины относительно другой и требуется найти первую величину или, говоря точнее, дана производная функции, а требуется найти саму функцию, т. е. по заданной функции найти одну из ее первообразных. Эта задача также решается с помощью определенного интеграла, так как такой первообразной является, например, определенный интеграл с переменным верхним пределом. В качестве примера подобной задачи рассмотрим вычисление длины дуги кривой.

Пусть кривая Γ задана параметрическим векторным представлением

$$\boldsymbol{r} = \boldsymbol{r}(t), \quad a \leqslant t \leqslant b,$$

где функция $\boldsymbol{r}(t)$ непрерывно дифференцируема на отрезке $[a, b]$. Тогда, как мы знаем, кривая Γ спрямляема и переменная длина дуги $s(t)$, отсчитываемая от начальной точки (ее радиусом-вектором служит $\boldsymbol{r}(a)$) кривой Γ является также непрерывно дифференцируемой функцией параметра t на отрезке $[a, b]$, причем (см. п. 16.5) $\frac{ds}{dt} = \left|\frac{dr}{dt}\right|$. Поэтому, в силу формулы Ньютона—Лейбница, замечая, что $s(a) = 0$, для длины $S = s(b)$ кривой Γ получаем

$$S = s(b) - s(a) = \int_a^b \frac{ds}{dt}\, dt,$$

откуда $S = \int_a^b \left|\frac{dr}{dt}\right| dt$.

Если $\boldsymbol{r}(t) = (x(t), y(t), z(t))$, то

$$S = \int_a^b \sqrt{x'^2(t) + y'^2(t) + z'^2(t)}\, dt. \qquad (28.29)$$

В том случае, когда кривая Γ является графиком непрерывно дифференцируемой функции $y = f(x)$, $a \leqslant x \leqslant b$, формула (28.29) принимает вид

$$S = \int_a^b \sqrt{1 + f'^2(x)}\, dx. \qquad (28.30)$$

Примеры. 1. Найдем длину S дуги параболы $y = ax^2$, $0 \leqslant x \leqslant b$. Замечая, что $y' = 2ax$, согласно формуле (28.30), имеем

$$S = \int_0^b \sqrt{1 + 4a^2x^2}\, dx. \qquad (28.31)$$

Неопределенный интеграл $I = \int \sqrt{1 + 4a^2x^2}\, dx$ вычислим следующим образом: проинтегрируем его сначала по частям; затем к числителю дроби, получившейся под знаком интеграла, прибавим и вычтем единицу, произведем деление и проинтегрируем (с помощью подстановки $y = 2ax$) получившуюся дробь:

$$I = \int \sqrt{1 + 4a^2x^2}\, dx = x\sqrt{1 + 4a^2x^2} - \int \frac{4a^2x^2}{\sqrt{1 + 4a^2x^2}}\, dx =$$

$$= x\sqrt{1 + 4a^2x^2} - \int \sqrt{1 + 4a^2x^2}\, dx + \int \frac{dx}{\sqrt{1 + 4a^2x^2}} =$$

$$= x\sqrt{1 + 4a^2x^2} - I + \frac{1}{2a} \ln \left|2ax + \sqrt{1 + 4a^2x^2}\right|.$$

Это равенство, рассматриваемое как уравнение относительно интеграла I, дает возможность найти его значение:

$$I = \frac{1}{2} x\sqrt{1 + 4a^2x^2} + \frac{1}{4a} \ln \left|2ax + \sqrt{1 + 4a^2x^2}\right| + C.$$

Теперь легко найти величину интеграла (28.31):

$$S = \frac{1}{2} b\sqrt{1 + 4a^2b^2} + \frac{1}{4a} \ln \left|2ab + \sqrt{1 + 4a^2b^2}\right|.$$

2. Найдем длину астроиды $x = a\cos^3 t$, $y = a\sin^3 t$ (см. рис. 91). Астроида симметрична относительно координатных осей. Ее части, лежащей в первой четверти, соответствует изменение параметра t от 0 до $\pi/2$. Вычислим длину S этой части (равной, очевидно, 1/4 длины всей астроиды). Заметив, что $x' = -3a\cos^2 t \sin t$, $y' = 3a\sin^2 t \cos t$, по формуле (28.29) (в которой следует положить $z' = 0$) получим

$$S = \int_0^{\pi/2} \sqrt{9a^2\cos^4 t\sin^2 t + 9a^2\sin^4 t\cos^2 t}\; dt =$$
$$= \frac{3a}{2}\int_0^{\pi/2} \sin 2t\, dt = \frac{3a}{2}.$$

Таким образом, длина всей астроиды равна $6a$.

3. Рассмотрим задачу о нахождении длины S дуги эллипса $x = a\sin t$, $y = b\cos t$, $0 \leqslant t \leqslant 2\pi$, $0 < b \leqslant a$, от верхнего конца малой полуоси до его точки, соответствующей значению параметра $t \in [0, 2\pi]$. Положим $\varepsilon = \sqrt{a^2 - b^2}/a$ (ε — эксцентриситет эллипса); тогда

$$\sqrt{x'^2 + y'^2} = \sqrt{a^2\cos^2 t + b^2\sin^2 t} = a\sqrt{1 - \varepsilon^2\sin^2 t},$$

поэтому

$$S = a\int_0^t \sqrt{1 - \varepsilon^2\sin^2 t}\; dt, \quad 0 \leqslant \varepsilon < 1. \qquad (28.32)$$

Получился эллиптический интеграл второго рода, который, как известно (см. п. 22.6), не выражается через элементарные функции, т. е. формула (28.32) в данном случае является окончательным ответом. Приближенные значения длин дуг эллипса можно получить либо непосредственно, вычислив приближенно интеграл (28.32), либо воспользовавшись имеющимися таблицами значений эллиптических интегралов.

УПРАЖНЕНИЯ. 1. Доказать, что если плоская кривая задана в полярных координатах непрерывно дифференцируемым представлением $r = r(\varphi)$, $\alpha \leqslant \varphi \leqslant \beta$, то для ее длины S справедлива формула

$$S = \int_\alpha^\beta \sqrt{r^2 + r'^2}\, d\varphi. \qquad (28.33)$$

2. Найти длину дуги логарифмической спирали $r = ae^{b\varphi}$ от точки (φ_0, r_0) до точки (φ, r).

Интегральная формула для длины кривой позволяет выразить ее длину не только как верхнюю грань длин всевозможных вписанных в нее ломаных, но и как их предел длин вписанных в нее ломаных при условии, что мелкости соответствующих разбиений стремятся к нулю. Чтобы это доказать, потребуется следующая лемма.

ЛЕММА. *Пусть* $\Gamma = \{\boldsymbol{r} = \boldsymbol{r}(s);\, 0 \leqslant s \leqslant S\}$ *— непрерывно дифференцируемая кривая в* $\boldsymbol{R}^3$, s *— ее переменная длина дуги и* $\Delta \boldsymbol{r} = \boldsymbol{r}(s + \Delta s) - \boldsymbol{r}(s)$. *Тогда отношение* $\frac{|\Delta r|}{|\Delta s|}$ *стремится к единице при* $\Delta s \to 0$ *равномерно на отрезке* $[0, S]$. *Это означает, что для любого* $\varepsilon > 0$ *существует такое* $\delta > 0$, *что для любой точки* $s \in [0, S]$ *и для любого приращения* Δs $(s + \Delta s \in [0, S])$, *удовлетворяющего условию* $|\Delta s| < \delta$, *выполняется неравенство*

$$\left|\frac{|\Delta r|}{|\Delta s|} - 1\right| < \varepsilon.$$

Доказательство. Предварительно заметим, что для любого числа $a \geqslant 0$ справедливо неравенство $|a - 1| \leqslant |a^2 - 1|$, так как в силу очевидного неравенства $a + 1 \geqslant 1$ имеем

$$|a - 1| = \left|\frac{a^2 - 1}{a + 1}\right| \leqslant |a^2 - 1|.$$

Пусть $r(s) = (x(s), y(s), z(s))$. В силу непрерывной дифференцируемости кривой производные $x'(s)$, $y'(s)$, $z'(s)$ ограничены и равномерно непрерывны на отрезке $[0, S]$. Поэтому, во-первых, существует такое число $c > 0$, что для всех $s \in [0, S]$ выполняются неравенства

$$\left|\frac{dx(s)}{ds}\right| \leqslant c, \quad \left|\frac{dy(s)}{ds}\right| \leqslant c, \quad \left|\frac{dz(s)}{ds}\right| \leqslant c,$$

а, во-вторых, для любого $\varepsilon > 0$ существует такое $\delta > 0$, что для всех точек $s \in [0, S]$ и всех приращений Δs, для которых $|\Delta s| < \delta$ и $s + \Delta s \in [0, S]$, выполняются неравенства:

$$\left|\frac{dx(s + \Delta s)}{ds} - \frac{dx(s)}{ds}\right| < \varepsilon, \quad \left|\frac{dy(s + \Delta s)}{ds} - \frac{dy(s)}{ds}\right| < \varepsilon,$$
$$\left|\frac{dz(s + \Delta s)}{ds} - \frac{dz(s)}{dz}\right| < \varepsilon.$$

Пусть $\varepsilon > 0$ задано и для него выбрано указанное $\delta > 0$, пусть $s \in [0, S]$, $s + \Delta s \in [0, S]$ и $|\Delta s| < \delta$. Заметив, что

$$\left(\frac{dx(s)}{ds}\right)^2 + \left(\frac{dy(s)}{ds}\right)^2 + \left(\frac{dz(s)}{ds}\right)^2 = \left|\frac{dr}{ds}\right|^2 = 1$$

и, применив формулу конечных приращений Лагранжа к функциям $x(s)$, $y(s)$, $z(s)$, получим:

$$\left|\left|\frac{\Delta \boldsymbol{r}}{\Delta s}\right| - 1\right| \leqslant \left|\left|\frac{\Delta \boldsymbol{r}}{ds}\right|^2 - 1\right| = \left|\left(\frac{x(s+\Delta s) - x(s)}{\Delta s}\right)^2 + \left(\frac{y(s+\Delta s) - y(s)}{\Delta s}\right)^2 + \right.$$

$$\left. + \left(\frac{z(s+\Delta s) - z(s)}{\Delta s}\right)^2 - 1\right| = \left|\left(\frac{dx(s+\theta_1 \Delta s)}{ds}\right)^2 - \left(\frac{dx(s)}{ds}\right)^2 + \right.$$

$$\left. + \left(\frac{dy(s+\theta_2 \Delta s)}{ds}\right)^2 - \left(\frac{dy(s)}{ds}\right)^2 + \left(\frac{dz(s+\theta_3 \Delta s)}{ds}\right)^2 - \left(\frac{dz(s)}{ds}\right)^2\right| \leqslant$$

$$\leqslant \left|\frac{dx(s+\theta_1 \Delta s)}{ds} - \frac{dx(s)}{ds}\right| \left(\left|\frac{dx(s+\theta_1 \Delta s)}{ds}\right| + \left|\frac{dx(s)}{ds}\right|\right) +$$

$$+ \left|\frac{dy(s+\theta_2 \Delta s)}{ds} - \frac{dy(s)}{ds}\right| \left(\left|\frac{dy(s+\theta_2 \Delta s)}{ds}\right| + \left|\frac{dy(s)}{ds}\right|\right) +$$

$$+ \left|\frac{dz(s+\theta_3 \Delta s)}{ds} - \frac{dz(s)}{ds}\right| \left(\left|\frac{dz(s+\theta_3 \Delta s)}{ds}\right| + \left|\frac{dz(s)}{ds}\right|\right) < 6c\varepsilon,$$

где $0 < \theta_1, \theta_2, \theta_3 < 1$. Лемма доказана. □

ТЕОРЕМА 3. *Пусть* $\Gamma = \{\boldsymbol{r} = \boldsymbol{r}(s);\ 0 \leqslant s \leqslant S\}$ *— непрерывно дифференцируемая кривая в* $\boldsymbol{R}^3$, s *— ее переменная длина дуги,* $\tau = \{s_i\}_{i=0}^{i=i_\tau}$ *— разбиение отрезка* $[0, S]$, $\lambda_\tau = \sum_{i=1}^{i_\tau} |\boldsymbol{r}(s_i) - \boldsymbol{r}(s_{i-1})|$; *тогда* $S = \lim\limits_{|\tau| \to 0} \lambda_\tau$.

Здесь λ_τ является, очевидно, длиной вписанной в кривую Γ ломаной с вершинами в точках $r(s_i)$, $i = 0, 1, 2, \ldots, i_\tau$.

Доказательство. Положим $\Delta s_i = s_i - s_{i-1}$, $\Delta \boldsymbol{r}_i = \boldsymbol{r}(s_i) - \boldsymbol{r}(s_{i-1})$, $i = 1, 2, \ldots, i_\tau$. Заметив, что $S = \sum_{i=1}^{i_\tau} \Delta s_i$ и $\lambda_\tau = \sum_{i=1}^{i_\tau} |\Delta \boldsymbol{r}_i|$, получим

$$|S - \lambda_\tau| = \left|\sum_{i=1}^{i_\tau} \Delta s_i - \sum_{i=1}^{i_\tau} |\Delta \boldsymbol{r}_i|\right| \leqslant \sum_{i=1}^{i_\tau} \left|1 - \left|\frac{\Delta \boldsymbol{r}_i}{\Delta s_i}\right|\right| \Delta s_i.$$

Согласно лемме, для любого $\varepsilon > 0$ существует такое $\delta > 0$, что как только $|\Delta s_i| < \delta$, то имеет место неравенство

$$\left|\left|\frac{\Delta \boldsymbol{r}_i}{\Delta s_i}\right| - 1\right| < \frac{\varepsilon}{S}.$$

Поэтому для всякого разбиения τ мелкости $|\tau| < \delta$ выполняется неравенство

$$|S - \lambda_\tau| < \frac{\varepsilon}{S} \sum_{i=1}^{i_\tau} \Delta s_i = \varepsilon.$$

Это и означает, что $\lim\limits_{|\tau| \to 0} \lambda_\tau = S$. □

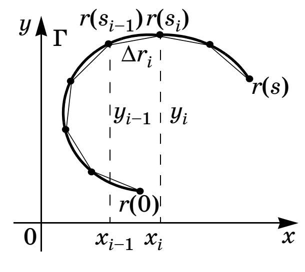


Рис. 137

28.5. Площадь поверхности вращения

Понятие поверхности и ее площади будет специально изучаться в § 50. Здесь же мы ограничимся специальным случаем поверхностей, образованных вращением кривых вокруг некоторых осей. Как всегда, будем предполагать, что в пространстве $\boldsymbol{R}^3$ фиксирована прямоугольная декартова система координат. Множество, получающееся в результате вращения плоской кривой вокруг некоторой прямой, лежащей в плоскости кривой, называется *поверхностью вращения*.

Пусть $\Gamma = \{\boldsymbol{r} = \boldsymbol{r}(t),\ a \leqslant t \leqslant b\}$ — кривая, лежащая в полуплоскости $y > 0$ плоскости переменных x, y, $\tau = \{t_i\}_{i=0}^{i=i_\tau}$ — разбиение отрезка $[a, b]$. Впишем в кривую Γ ломаную с вершинами в точках $r(t_i) = (x_i, y_i)$, $i = 0, 1, 2, \ldots, i_\tau$ (рис. 137). При вращении звена $\Delta \boldsymbol{r}_i = \boldsymbol{r}(t_i) - \boldsymbol{r}(t_{i-1})$ этой ломаной вокруг оси Ox получится поверхность усеченного конуса (в частности, быть может, цилиндра) с площадью

$$l_i = \pi(y_{i-1} + y_i)|\Delta \boldsymbol{r}_i|,$$

а при вращении всей ломаной — поверхность с площадью

$$L_\tau = \sum_{i=1}^{i_\tau} l_i = \pi \sum_{i=1}^{i_\tau} (y_{i-1} + y_i)|\Delta \boldsymbol{r}_i|.$$

Определение 1. *Если существует предел* $\lim\limits_{|\tau| \to 0} L_\tau$, *то он называется площадью* L *поверхности, образованной вращением кривой* Γ *вокруг оси* Ox.

Таким образом,

$$L \overset{\text{def}}{=} \lim_{|\tau| \to 0} L_\tau. \tag{28.34}$$

ТЕОРЕМА 4. *Пусть* $\Gamma = \{\boldsymbol{r} = \boldsymbol{r}(t);\ a \leqslant t \leqslant b\}$ — *непрерывно дифференцируемая кривая без особых точек, лежащая в*

полуплоскости $y > 0$ плоскости переменных x, y. Тогда для площади L поверхности, полученной вращением кривой Γ *вокруг оси Ox, справедлива формула*

$$L = 2\pi\int_a^b y\sqrt{x_t'^2 + y_t'^2}\,dt = 2\pi\int_0^S y(s)\,ds, \qquad (28.35)$$

где s — переменная длина дуги кривой Γ, $0 \leqslant s \leqslant S$.

Доказательство. Как известно, при сделанных в этой теореме предположениях (см. п. 16.5) функция $s = s(t)$, $a \leqslant t \leqslant b$, является допустимым преобразованием параметра, и, следовательно, длина дуги s может быть принята за параметр:

$$\Gamma = \{\boldsymbol{r} = \boldsymbol{r}(s) = (x(s), y(s));\ 0 \leqslant s \leqslant S\}.$$

Пусть $\tau = \{s_i\}_{i=0}^{i=i_\tau}$ — разбиение отрезка $[0, S]$,

$$\Delta\boldsymbol{r}_i = \boldsymbol{r}(s_i) - \boldsymbol{r}(s_{i-1}),\quad \Delta s_i = s_i - s_{i-1},\quad i = 1, 2, \dots, i_\tau.$$

Сравним сумму

$$L_\tau = \pi\sum_{i=1}^{i_\tau}(y_{i-1} + y_i)|\Delta\boldsymbol{r}_i|,\ y_i \overset{\text{def}}{=} y(s_i),\ i = 0, 1, \dots, i_\tau, \qquad (28.36)$$

с интегральной суммой (функции $2\pi y(s)$)

$$\sigma_\tau = 2\pi\sum_{i=1}^{i_\tau} y_i\Delta s_i. \qquad (28.37)$$

Для этого заметим, что функция $y(s)$, будучи непрерывной на отрезке $[0, S]$, ограничена на нем, т. е. существует такая постоянная $M > 0$, что для всех $s \in [0, S]$ выполняется неравенство $|y(s)| \leqslant M$. Обозначая через $\omega(\delta; y)$ модуль непрерывности функции $y(s)$, а через λ_τ — длину ломаной с вершинами в точках $\boldsymbol{r}(s_i)$ и заметив, что $|\Delta\boldsymbol{r}_i| \leqslant \Delta s_i$, $i = 1, 2, \dots, i_\tau$, получим

$$|\sigma_\tau - L_\tau| = \left|\sum_{i=1}^{i_\tau}\ \pi\sum_{i=1}^{i_\tau} 2y_i\Delta s_i - \pi\sum_{i=1}^{i_\tau}[2y_i + (y_{i-1} - y_i)]\,|\Delta\boldsymbol{r}_i|\right| \leqslant$$

$$\leqslant 2\pi\sum_{i=1}^{i_\tau}|y_i|(\Delta s_i - |\Delta\boldsymbol{r}_i|) + \pi\sum_{i=1}^{i_\tau}|y_i - y_{i-1}|\,|\Delta\boldsymbol{r}_i| \leqslant$$

$$\leqslant 2\pi M\left|\ \sum_{i=1}^{i_\tau}\Delta s_i - \sum_{i=1}^{i_\tau}|\Delta\boldsymbol{r}_i|\ \right| + \pi\omega(|\tau|; y)\sum_{i=1}^{i_\tau} + |\Delta\boldsymbol{r}_i| =$$

$$= 2\pi M(S - \lambda_\tau) + \pi\,\omega(|\tau|; y)\,\lambda_\tau.$$

Здесь $\lim\limits_{|\tau|\to 0}(S - \lambda_\tau) = 0$ (см. теорему 3), $\lim\limits_{|\tau|\to 0}\omega(|\tau|, y) = 0$ (см. теорему 5 в п. 19.7) и $0 \leqslant \lambda_\tau \leqslant S$. Поэтому $\lim\limits_{|\tau|\to 0}(\sigma_\tau - L_\tau) = 0$,

а так как $\lim\limits_{|\tau| \to 0} \sigma_\tau = 2\pi\int\limits_0^S y(s)\,ds$, то и $\lim\limits_{|\tau| \to 0} L_\tau = 2\pi\int\limits_0^S y(s)\,ds$. Сделав в последнем интеграле замену переменного $s = s(t)$ и вспоминая, что $ds = \sqrt{x'^2 + y'^2}\,dt$, получим

$$L = 2\pi\int\limits_a^b y\sqrt{x'^2 + y'^2}\,dt. \ \square$$

Если кривая Γ задана явным уравнением $y = f(x)$, $a \leqslant x \leqslant b$, то формула для площади поверхности, образованной вращением графика функции f вокруг оси Ox, имеет вид

$$L = 2\pi\int\limits_a^b y\sqrt{1 + y'^2}\,dx. \qquad (28.38)$$

Предложенный вывод формулы (28.35) имеет тот недостаток, что в нем уже использовалось понятие площади поверхности и ее аддитивность, правда, лишь в простейшем случае — для поверхностей усеченного конуса и их объединений. Можно ввести общее понятие площади поверхности, не используя понятие площади поверхности для каких-либо элементарных поверхностей, и получить ее необходимые свойства. Эти вопросы будут рассмотрены в п. 50.9.

Примеры. 1. Найдем площадь S сферы радиуса r. Указанная сфера может быть получена вращением полуокружности $y = \sqrt{r^2 - x^2}$, $-r \leqslant x \leqslant r$, вокруг оси Ox. Однако это явное представление полуокружности не является непрерывно дифференцируемым: производная $y' = -\dfrac{x}{\sqrt{r^2 - x^2}}$ обращается в бесконечность при $x = \pm r$. Гораздо удобнее взять параметрическое представление полуокружности $x = r\cos t$, $y = r\sin t$, $0 \leqslant t \leqslant \pi$. Тогда $x' = -r\sin t$, $y' = r\cos t$; поэтому площадь S поверхности сферы радиуса r легко вычисляется по формуле (28.35):

$$S = \int\limits_0^\pi y\sqrt{x'^2 + y'^2}\,dt = 2\pi r^2\int\limits_0^\pi \sin t\,dt = 4\pi r^2.$$

2. Найдем площадь S поверхности, образованной вращением вокруг оси Ox дуги цепной линии (см. рис. 136)

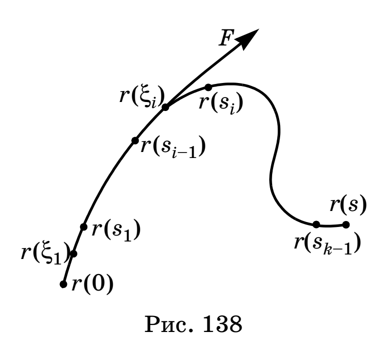


Рис. 138

$y = a\operatorname{ch}\dfrac{x}{a}$, $-b \leqslant x \leqslant b$ (эта поверхность называется *катеноидом*).

По формуле (28.38) имеем

$$S = 2\pi a\int_{-b}^{b}\operatorname{ch}\frac{x}{a}\sqrt{1+\operatorname{sh}^2\frac{x}{a}}\,dx =$$

$$= 2\pi a\int_{-b}^{b}\operatorname{ch}^2\frac{x}{a}\,dx = \pi a\int_{-b}^{b}\left(1+\operatorname{ch}\frac{2x}{a}\right)dx =$$

$$= \pi a\left(2b + a\operatorname{sh}\frac{2b}{a}\right).$$

28.6. Работа силы

Пусть материальная точка M движется по непрерывно дифференцируемой кривой $\Gamma = \{r = r(s)\}$, где s — переменная длина дуги, $0 \leqslant s \leqslant S$. Пусть на рассматриваемую материальную точку, находящуюся в положении $r(s)$, действует сила $F(s)$, направленная по касательной к траектории в направлении движения (точнее, $F(s)$ — численная величина этой силы).

Возьмем какое-либо разбиение $\tau = \{s_i\}_{i=0}^{i=i_\tau}$ отрезка $[0, S]$. Ему соответствует разбиение траектории Γ на части

$$\Gamma_i = \{r(s),\ s_{i-1} \leqslant s \leqslant s_i\},\quad i = 1, \ldots, i_\tau.$$

Выберем произвольно по точке $\xi_i \leqslant [s_{i-1}, s_i]$ (рис. 138). Величина $F(\xi_i)\Delta s_i$, $\Delta s_i = s_i - s_{i-1}$, $i = 1, 2, \ldots, i_\tau$, называется *элементарной работой* силы F на участке Γ_i и принимается за приближенное значение работы, которую производит сила F, воздействующая на материальную точку, когда последняя проходит кривую Γ_i. Сумма всех элементарных работ $\sum\limits_{i=1}^{i_\tau} F(\xi_i)\Delta s_i$ является интегральной суммой Римана функции $F(s)$.

Определение 2. *Предел, к которому стремится сумма* $\sum\limits_{i=1}^{i_\tau} F(\xi_i)\Delta s_i$ *всех элементарных работ, когда мелкость* $|\tau|$ *разбиения* τ *стремится к нулю, называется работой силы* F *вдоль кривой* Γ.

Таким образом, если обозначить эту работу буквой W, то, в силу данного определения,

$$W = \lim_{|\tau| \to 0} \sum_{i=1}^{i_\tau} F(\xi_i)\Delta s_i,$$

следовательно,

$$W = \int_0^S F(s)\, ds. \tag{28.39}$$

Если положение точки на траектории ее движения описывается с помощью какого-либо другого параметра t (например, времени) и если величина пройденного пути $s = s(t)$, $a \leqslant t \leqslant b$ является непрерывно дифференцируемой функцией, то из формулы (28.39) получим

$$W = \int_a^b F[s(t)]s'(t)\, dt.$$

28.7. Вычисление статических моментов и координат центра тяжести кривой

Пусть M — материальная точка массы m с координатами x и y. Произведения my и mx называются ее *статическими моментами* соответственно *относительно осей Ox и Oy*.

Пусть $\Gamma = \{r(s), 0 \leqslant s \leqslant S\}$ — спрямляемая кривая, лежащая в полуплоскости $y > 0$, s — переменная длина дуги, кривая Γ имеет массу и масса ее дуги прямо пропорциональна длине дуги; если Δm — масса дуги длиной Δs, то $\Delta m = \rho\Delta s$, где ρ — некоторая постоянная, называемая *линейной плотностью кривой* Γ. Такие кривые в механике называются *однородными*. Поскольку $\rho = \frac{\Delta m}{\Delta s}$, плотность равна массе длины дуги кривой, приходящейся на единицу длины дуги. Будем считать для простоты, что $\rho = 1$, т. е. что масса части кривой длины Δs также равна Δs, в частности, что масса всей кривой численно равна S.

Пусть теперь $\tau = \{s_i\}_{i=0}^{i=i_\tau}$ — какое-либо разбиение отрезка $[0, S]$, $\Delta s_i = s_i - s_{i-1}$, $i = 1, 2, \ldots, i_\tau$. Разбиению τ соответствует разбиение кривой Γ на части $\Gamma_i = \{r(s), s_{i-1} \leqslant s \leqslant s_i\}$. Выберем по какой-либо точке $\xi_i \in [s_{i-1}, s_i]$ и положим $x_i = x(\xi_i)$, $y_i = y(\xi_i)$, $i = 1, 2, \ldots, i_\tau$.

Величины $y_i\Delta s_i$ при любом выборе указанных точек ξ_i называются *элементарными статическими моментами* час-

ти Γ_i кривой Γ относительно оси Ox. Очевидно, элементарный статический момент Γ_i численно равен статическому моменту материальной точки массы Δs_i с ординатой y_i, т. е. мы как бы заменили данную непрерывную кривую Γ отдельными i_τ материальными точками.

Определение 3. *Предел, к которому стремится сумма*

$$\sum_{i=1}^{i_\tau} y_i \Delta s_i \tag{28.40}$$

всех элементарных статических моментов, когда мелкость разбиения τ стремится к нулю, называется статическим моментом M_x кривой Γ относительно оси Ox.

Этот предел всегда существует, так как, по определению кривой, функция $r = r(s)$, а значит, и координатные функции $x = x(s)$, $y = y(s)$ непрерывны на отрезке $[0, S]$; сумма же (28.39) является интегральной суммой Римана функции $y(s)$ и поэтому при $|\tau| \to 0$ стремится к интегралу $\int_0^S y(s)\, ds$. Таким образом,

$$M_x = \int_0^S y\, ds. \tag{28.41}$$

Аналогично определяется и вычисляется статический момент M_y кривой Γ относительно оси Ox:

$$M_y = \int_0^S x\, ds. \tag{28.42}$$

Определение 4. *Точка плоскости $P = (x_0, y_0)$, обладающая тем свойством, что если в нее поместить материальную точку массы, равной массе кривой (в рассматриваемом случае массы S), то эта точка относительно координатных осей имеет статические моменты, численно равные статическим моментам кривой относительно тех же осей, называется центром тяжести данной кривой.*

Таким образом, $Sx_0 = M_y$, $Sy_0 = M_x$, откуда, в силу формул (28.41) и (28.42), для координат центра тяжести получаем формулы

$$x_0 = \frac{1}{S}\int_0^S x\, ds, \quad y_0 = \frac{1}{S}\int_0^S y\, ds. \tag{28.43}$$

Сравнивая формулы для ординаты центра тяжести кривой $y_0 S = \int_0^S y\, ds$ и для площади L поверхности, полученной от вра-

щения этой кривой вокруг оси Ox: $L = 2\pi\int\limits_0^S y\,ds$, получим интересное соотношение $L = 2\pi y_0 S$ (здесь под кривой понимается непрерывно дифференцируемая кривая без особых точек, лежащая в полуплоскости $y > 0$), которое составляет содержание так называемой п е р в о й т е о р е м ы Г у л ь д и н а[1].

ТЕОРЕМА 5 (первая теорема Гульдина). *Площадь поверхности, полученной от вращения кривой около некоторой не пересекающей ее оси, равна длине этой кривой, умноженной на длину окружности, описанной центром тяжести этой кривой.*

В том случае, когда известно положение центра тяжести кривой, теорема Гульдина позволяет просто находить площадь соответствующей поверхности вращения. Например, площадь поверхности, полученной от вращения окружности $(x - a)^2 + y^2 = r^2$, $0 < r < a$, вокруг оси Oy (такая поверхность называется *тором*), легко вычисляется указанным способом: $L = 2\pi a \cdot 2\pi r = 4\pi^2 ar$, так как центр тяжести окружности совпадает с ее центром.

В качестве примера вычисления центра тяжести кривой по формуле (28.43) найдем центр тяжести цепной линии $y = a\,\mathrm{ch}\,\frac{x}{a}$, $-b \leqslant x \leqslant b$. В силу симметрии цепной линии относительно оси Oy, имеем $M_y = 0$. Действительно, выбирая за начало отсчета дуг точку цепной линии, лежащую на оси Oy, и обозначив длину всей цепной линии через $2S$, получим

$$M_y = \int\limits_{-S}^{S} x(s)\,ds = 0,$$

так как $x(s)$ — нечетная функция. Из равенства $M_y = 0$, в силу формулы (28.43), следует, что $x_0 = 0$.

В силу теоремы Гульдина, $L_x = 2\pi y_0 \cdot 2S$, где $2S$ — длина кривой, в данном случае рассматриваемой цепной линии, а L_x — площадь поверхности вращения, образованной вращением этой линии вокруг оси Ox. Площадь L_x была вычислена в п. 28.5:

$$L_x = \pi a\left(2b + a\,\mathrm{sh}\,\frac{2b}{a}\right),$$

[1] П. Гульдин (1577—1643) — швейцарский математик.

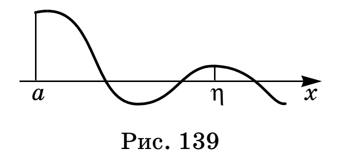


Рис. 139

а длина $2S$ цепной линии легко вычисляется по формуле (28.30):

$$2S = \int_{-b}^{b} \sqrt{1 + y'^2}\, dx = \int_{-b}^{b} \sqrt{1 + \operatorname{sh}^2 \frac{x}{a}}\, dx =$$

$$= \int_{-b}^{b} \operatorname{ch} \frac{x}{a}\, dx = a \operatorname{sh} \frac{x}{a} \bigg|_{-b}^{b} = 2a \operatorname{sh} \frac{b}{a};$$

в силу формулы (28.43), получим

$$y_0 = \frac{2b + a\operatorname{sh}\frac{2b}{a}}{4\operatorname{sh}\frac{b}{a}}.$$

УПРАЖНЕНИЕ 3. Доказать единственность центра тяжести непрерывно дифференцируемой кривой, иначе говоря, что точка плоскости, определяемая формулами (28.43), не зависит от выбора декартовых координат на плоскости.

§ 29.

Несобственные интегралы

29.1. Определение несобственных интегралов

Функция, не ограниченная на отрезке, не интегрируема на нем по Риману (теорема 1 п. 23.3). Если функция определена на бесконечном промежутке, то нельзя говорить об ее интегрируемости по Риману просто потому, что определение интеграла относится только к функциям, заданным на отрезке. В настоящем параграфе понятие интеграла обобщается как на случай функций, определенных на неограниченных промежутках, так и на случай функций, определенных на ограниченных промежутках, но не ограниченных на них. Это делается с помощью предельного перехода, дополнительного к пределу, с помощью которого вводится интеграл Римана.

Пусть функция f определена на конечном или бесконечном промежутке $[a, b)$, $-\infty < a < b \leqslant +\infty$, и для любого числа $\eta \in [a, b)$ она интегрируема на отрезке $[a, \eta]$ (рис. 139).

Определение 1. *Функция* $F(\eta) = \int\limits_a^\eta f(x)\,dx$ *верхнего предела интегрирования,* $a \leqslant \eta < b$*, называется несобственным интегралом и обозначается* $\int\limits_a^b f(x)\,dx$.

Если существует конечный предел $\lim\limits_{\eta \to b}\int\limits_a^\eta f(x)\,dx$*, то несобственный интеграл* $\int\limits_a^b f(x)\,dx$ *называется сходящимся, а если этот предел не существует, то — расходящимся.*

В случае, когда несобственный интеграл сходится, говорят также, что он существует, а если расходится, то — не существует.

Если интеграл $\int\limits_a^b f(x)\,dx$ сходится, то предел $\lim\limits_{\eta \to b}\int\limits_a^\eta f(x)\,dx$ обозначают тем же символом, что и сам интеграл[1], т. е.

$$\int\limits_a^b f(x)\,dx = \lim_{\eta \to b}\int\limits_a^\eta f(x)\,dx \qquad (29.1)$$

и для краткости обычно также называют несобственным интегралом (или, подробнее, значением несобственного интеграла).

В отличие от несобственного интеграла обычный интеграл Римана называют иногда *собственным интегралом*.

Существование несобственного интеграла $\int\limits_a^b f(x)\,dx$ эквивалентно существованию несобственного интеграла $\int\limits_c^b f(x)\,dx$ при любом $c \in (a, b)$. В самом деле, интеграл $\int\limits_a^\eta f(x)\,dx$ отличается от интеграла $\int\limits_c^\eta f(x)\,dx$ (при $c < \eta < b$) на конечную, не зависящую от η величину $\int\limits_a^c f(x)\,dx$:

$$\int\limits_a^\eta f(x)\,dx = \int\limits_a^c f(x)\,dx + \int\limits_c^\eta f(x)\,dx.$$

[1] С подобной ситуацией мы встречались и раньше, например обозначая одним и тем же символом функцию x^2 и значение этой функции при значении аргумента, равном x.

Поэтому при $\eta \to b$ оба интеграла $\int\limits_a^\eta$ и $\int\limits_c^\eta$ одновременно имеют или не имеют предел, причем в случае его существования

$$\int\limits_c^b f(x)\,dx = \int\limits_a^c f(x)\,dx + \int\limits_c^b f(x)\,dx. \qquad (29.2)$$

Из определения (29.1) несобственного интеграла и из (29.2) следует, что если интеграл $\int\limits_c^b f(x)\,dx$ сходится, то

$$\lim_{c \to b} \int\limits_c^b f(x)\,dx = 0. \qquad (29.3)$$

Отметим, что выполнение этого условия нельзя принять в качестве определения сходящегося интеграла $\int\limits_a^b f(x)\,dx$, так как интеграл $\int\limits_c^b f(x)\,dx$ также является несобственным и говорить о его стремлении к нулю при $c \to b$ можно лишь уже имея определение сходящегося несобственного интеграла.

Подчеркнем, что определение (29.1) несобственного интеграла $\int\limits_a^b f(x)\,dx$ в случае конечного промежутка $[a, b)$ содержательно лишь когда функция f не ограничена в любой окрестности точки $x = b$, т. е. на любом интервале $(b - \varepsilon, b)$ $(0 < \varepsilon < b - a)$. Это связано с тем, что (см. замечание 2 в п. 23.7*) всякая функция, интегрируемая по Риману на любом отрезке $[a, \eta]$, $a \leqslant \eta < b < +\infty$, и ограниченная на полуинтервале $[a, b)$, является интегрируемой по Риману и на отрезке $[a, b]$ при любом ее доопределении в точке $x = b$. При этом интеграл Римана от таким образом доопределенной функции равен пределу (29.1) и тем самым не зависит от выбора дополнительного значения функции при $x = b$. В этом смысле интеграл Римана является частным случаем несобственного интеграла и можно говорить об интеграле Римана по конечному полуинтервалу $[a, b)$ от функции, заданной на этом промежутке (а в дальнейшем и об интеграле Римана по конечному интервалу).

В силу сказанного, теория несобственных интегралов содержательна, т. е. приводит к принципиально новым результатам лишь когда функция определена на бесконечном промежутке или конечном, причем в последнем случае она не ограничена

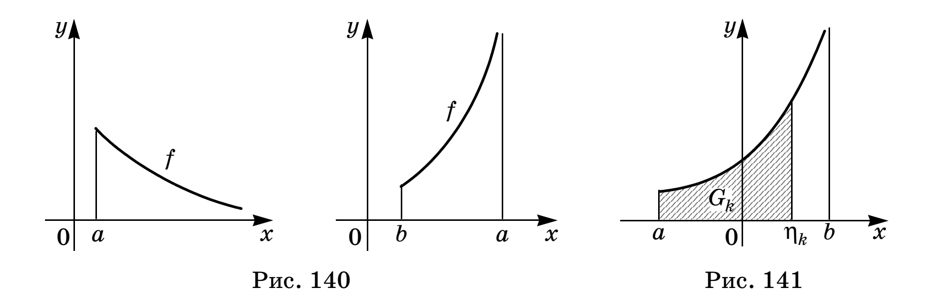


Рис. 140 Рис. 141

(рис. 140). Содержательность здесь понимается в том смысле, что для ограниченных подынтегральных функций, определенных на ограниченных промежутках, доказываемые ниже теоремы по существу доказаны раньше.

Если функция f неотрицательна и непрерывна на промежутке $[a, b)$, то несобственный интеграл $\int_a^b f(x)\,dx$, когда он существует, равен площади открытого множества

$$G = \{(x, y) : a < x < b;\ 0 < y < f(x)\},$$

т. е.

$$\int_a^b f(x)\,dx = \text{mes}\,G. \qquad (29.4)$$

Если $b = +\infty$ или если $b < -\infty$, но функция f не интегрируема по Риману на конечном полуинтервале $[a, b)$, то множество G неограниченно (на рис. 141 изображен случай конечного b). Для доказательства равенства (29.4) выберем какую-либо последовательность $\eta_k \in [a, b)$, $k = 1, 2, \ldots$, так, чтобы $\lim\limits_{k \to \infty} \eta_k = b$, и положим

$$G_k = \{(x, y) : a < x < \eta_k,\ 0 < y < f(x)\}.$$

Тогда, согласно теореме 1 из п. 28.1,

$$\text{mes}\,G_k = \int_a^{\eta_k} f(x)\,dx. \qquad (29.5)$$

Поскольку G_k — открытые множества, $k = 1, 2, \ldots$, и

$$G_1 \subset G_2 \subset \ldots \subset G_k \ldots \text{ и } \bigcup_{k=1}^{\infty} G_k = G,$$

в силу теоремы 2 п. 27.2,

$$\lim_{k \to \infty} \text{mes}\,G_k = \text{mes}\,G.$$

Согласно же определению несобственного интеграла,

$$\lim_{k \to \infty} \int_a^{\eta_k} f(x)\, dx = \int_a^b f(x)\, dx.$$

Поэтому, перейдя к пределу в равенстве (29.5) при $k \to \infty$, получим (39.4).

УПРАЖНЕНИЕ 1. Если функция f неотрицательна при $x \geqslant a$, и интеграл $\int\limits_a^{+\infty} f(x)\, dx$ сходится, то:

а) обязательно ли функция f ограничена;

б) обязательно ли функция f стремится к нулю при $x \to +\infty$;

в) каков будет ответ на эти вопросы, если дополнительно потребовать, что функция f была непрерывна при $x \geqslant a$? что она положительна?

Если функция f определена на полуинтервале вида $(a, b]$, $-\infty \leqslant a < b < +\infty$, и интегрируема по Риману на всех отрезках $[\xi, b]$, $a < \xi \leqslant b$, то несобственный интеграл $\int\limits_a^b f(x)\, dx$, его сходимость и расходимость определяются аналогично случаю полуинтервала $[a, b)$, а значение — по формуле

$$\int_a^b f(x)\, dx \overset{\text{def}}{=} \lim_{\xi \to a} \int_\xi^b f(x)\, dx. \qquad (29.6)$$

Если же функция f определена на интервале (a, b), $-\infty \leqslant a < b \leqslant +\infty$, и при некотором выборе точки $c \in (a, b)$ существуют несобственные интегралы $\int\limits_a^c f(x)\, dx$ (в смысле (29.6)) и $\int\limits_c^b f(x)\, dx$ (в смысле (29.1)), то, по определению, полагается

$$\int_a^b f(x)\, dx \overset{\text{def}}{=} \int_a^c f(x)\, dx + \int_c^b f(x)\, dx. \qquad (29.7)$$

При этом существование и значение интеграла $\int\limits_a^b f(x)\, dx$ не зависит от выбора точки $c \in (a, b)$. В самом деле, в рассматриваемом случае функция f, очевидно, интегрируема по Риману на любом отрезке $[\xi, \eta]$, $a < \xi < \eta < b$, и определение (29.7), в силу определений (29.1) и (29.6), равносильно следующему:

$$\int_a^b f(x)\, dx \overset{\text{def}}{=} \lim_{\substack{\xi \to a \\ \eta \to b}} \int_\xi^\eta f(x)\, dx, \quad a < \xi < \eta < b.$$

Здесь правая часть является пределом функции двух переменных ξ и η. Образно говоря, переменные ξ и η стремятся соответственно к a и b независимо друг от друга.

Если хотя бы один из интегралов $\int_a^c f(x)\,dx$ или $\int_c^b f(x)\,dx$ расходится, то говорят, что и интеграл $\int_a^b f(x)\,dx$ также расходится.

Определим теперь общее понятие несобственного интеграла от функции f по промежутку с концами a и b,

$$-\infty \leqslant a < b \leqslant +\infty.$$

Всякое множество точек $X = \{x_0, x_1, \ldots, x_k\}$ такое, что:

1) $a = x_0 < x_1 < \ldots < x_k = b$;

2) функция f интегрируема по Риману на любом отрезке, лежащем в рассматриваемом промежутке и не содержащем точек множества X, называется *правильным разбиением* этого промежутка относительно функции f (интервалы (x_0, x_1) и (x_{k-1}, x_k) могут оказаться неограниченными).

На каждом из интервалов (x_{i-1}, x_i), $i = 1, 2, \ldots, k$, имеет смысл рассматривать несобственный интеграл.

Если все интегралы $\int_{x_{i-1}}^{x_i} f(x)\,dx$, $i = 1, 2, \ldots, k$, существуют, то будем говорить, что существует (или, что то же самое, сходится) интеграл $\int_a^b f(x)\,dx$ и его значение определяется равенством

$$\int_a^b f(x)\,dx \stackrel{\text{def}}{=} \sum_{i=1}^{i_\tau} \int_{x_{i-1}}^{x_i} f(x)\,dx. \qquad (29.8)$$

Если хотя бы один из интегралов $\int_{x_{i-1}}^{x_i} f(x)\,dx$ расходится, то говорят, что «интеграл $\int_a^b f(x)\,dx$ *расходится*».

Из определений (29.7) и (29.8) следует, что несобственный интеграл в общем случае сводится к интегралам вида (29.1) и (29.6). Поэтому в дальнейшем ограничимся лишь изучением несобственных интегралов двух указанных видов.

УПРАЖНЕНИЯ. 2. Доказать, что существование и значение несобственного интеграла $\int_a^b f(x)\,dx$ в определении (29.8) не зависит от выбора точек x_i, $i = 0, 1, 2, \ldots, k$, удовлетворяющих сформулированным выше условиям.

3. Доказать, что если интеграл $\int\limits_a^{+\infty} f(x)\,dx$ сходится и существует $\lim\limits_{x\to+\infty} f(x)$, то этот предел равен нулю: $\lim\limits_{x\to+\infty} f(x) = 0$.

Примеры. 1. Покажем, что несобственный интеграл от функции $f(x) = 1/x$ по полуинтервалу $(0, 1]$ расходится. Действительно,

$$\int\limits_0^1 \frac{dx}{x} = \lim_{\xi\to+0}\int\limits_\xi^1 \frac{dx}{x} = \lim_{\xi\to+0} \ln x\Big|_\xi^1 = -\lim_{\xi\to+0} \ln\xi = +\infty.$$

Обычно проведенные вычисления записываются короче:

$$\int\limits_0^1 \frac{dx}{x} = \ln x\Big|_0^1 = +\infty.$$

2. Выясним, при каких $\alpha \neq 1$ сходится, а при каких — расходится интеграл от функции $f(x) = 1/x^\alpha$ по промежутку $(0, 1]$. Имеем

$$\int\limits_0^1 \frac{dx}{x^\alpha} = \frac{x^{1-\alpha}}{1-\alpha}\Big|_0^1 = \begin{cases} \dfrac{1}{1-\alpha} & \text{при } \alpha < 1, \\ +\infty & \text{при } \alpha > 1. \end{cases}$$

Отметим, что при $\alpha \leqslant 0$ рассматриваемый интеграл является собственным. Объединив результаты, полученные в примерах 1 и 2, получим

$$\int\limits_0^1 \frac{dx}{x^\alpha} \begin{cases} \text{сходится при } \alpha < 1, \\ \text{расходится при } \alpha \geqslant 1. \end{cases} \tag{29.9}$$

3. Рассмотрим теперь функцию $f(x) = 1/x^\alpha$ на бесконечном промежутке $[1, +\infty)$. Если $\alpha = 1$, то

$$\int\limits_1^{+\infty} \frac{dx}{x} = \ln x\Big|_1^{+\infty} = +\infty.$$

Если же $\alpha \neq 1$, то

$$\int\limits_1^{+\infty} \frac{dx}{x^\alpha} = \frac{x^{1-\alpha}}{1-\alpha}\Big|_1^{+\infty} = \begin{cases} \dfrac{1}{\alpha-1} & \text{при } \alpha > 1, \\ +\infty & \text{при } \alpha < 1. \end{cases}$$

Таким образом,

$$\int\limits_1^{+\infty} \frac{dx}{x^\alpha} \begin{cases} \text{сходится при } \alpha > 1, \\ \text{расходится при } \alpha \geqslant 1. \end{cases} \tag{29.10}$$

4. Если интеграл $\int\limits_a^b f(x)\,dx$ сходится, $0 \leqslant a < b \leqslant +\infty$, и

$$f^*(x) \overset{\text{def}}{=} f(-x),\ -b \leqslant x \leqslant -a,$$

то интеграл $\int\limits_{-b}^{-a} f^*(x)\,dx$ также сходится и

$$\int\limits_{-b}^{-a} f^*(x)\,dx = \int\limits_a^b f(x)\,dx.$$

Действительно, пусть, например, функция f интегрируема по Риману на любом отрезке $[a, \eta]$, $a \leqslant \eta < b$, и, следовательно,

$$\int\limits_a^b f(x)\,dx = \lim_{\eta \to b} \int\limits_a^{\eta} f(x)\,dx$$

(общий случай несобственного интеграла, в силу определения (29.8), сводится к подобным интегралам, точнее к интегралам вида (29.1) и (29.6)). В силу формулы (24.31) (см. п. 24.1), имеет место равенство

$$\int\limits_a^{\eta} f(x)\,dx = \int\limits_{-a}^{-\eta} f^*(x)\,dx.$$

Поэтому

$$\lim_{-\eta \to -b} \int\limits_{-a}^{-\eta} f^*(x)\,dx = \lim_{\eta \to b} \int\limits_a^{\eta} f(x)\,dx = \int\limits_a^b f(x)\,dx,$$

но предел левой части равенства и является несобственным интегралом $\int\limits_{-b}^{-a} f^*(x)\,dx$; таким образом,

$$\int\limits_{-b}^{-a} f^*(x)\,dx = \int\limits_a^b f(x)\,dx.$$

В частности, если функция f четная на отрезке $[-a, a]$ и интеграл $\int\limits_0^a f(x)\,dx$ сходится, то сходится и интеграл $\int\limits_{-a}^0 f(x)\,dx$, причем

$$\int\limits_{-a}^0 f(x)\,dx = \int\limits_0^a f(x)\,dx$$

и, следовательно,

$$\int\limits_{-a}^a f(x)\,dx = 2\int\limits_0^a f(x)\,dx.$$

5. Пусть функция f имеет период $T > 0$ и для некоторого числа a интеграл $\int\limits_a^{a+T} f(x)\,dx$ сходится; тогда для любого числа $b \geqslant a$ интеграл $\int\limits_b^{b+T} f(x)\,dx$ также сходится и

$$\int\limits_b^{b+T} f(x)\,dx = \int\limits_a^{a+T} f(x)\,dx.$$

Для доказательства достаточно заметить, что формулы (24.34) (см. п. 24.1) справедливы и в том случае, когда входящие в них интегралы несобственные. Это доказывается предельным переходом из равенства соответствующих собственных интегралов (их равенство следует из формулы (24.33)), пределом которых являются рассматриваемые несобственные интегралы.

Мы ввели новое понятие — понятие несобственного интеграла. Прежде всего естественно выяснить, какими свойствами обладает этот интеграл. Сохраняются ли для него свойства обычного интеграла? Возникают ли для несобственного интеграла (а если да, то какие) новые задачи и вопросы, специфические именно для него? Ответы на эти вопросы будут даны в дальнейших пунктах этого параграфа.

29.2. Формулы интегрального исчисления для несобственных интегралов

В силу свойств предела и определения значения несобственного интеграла как предела обычного интеграла Римана, на несобственные интегралы переносятся многие свойства определенного интеграла. Рассмотрим некоторые из них.

В этом и в дальнейших пунктах при рассмотрении свойств несобственных интегралов будем останавливаться более подробно лишь на интегралах от функций, определенных на конечных или бесконечных промежутках вида $[a, b)$ и интегрируемых по Риману на всех отрезках $[a, \eta]$, $a \leqslant \eta < b \leqslant +\infty$. Любые другие предположения будут специально оговариваться.

1^0 (формула Ньютона—Лейбница). *Если F — какая-либо первообразная функции f на полуинтервале* $[a, b)$, *то*

$$\int_a^b f(x)\,dx = F(b) - F(a) \equiv$$

$$\equiv \begin{cases} F(b-0) - F(a), & \text{если } b \text{ конечно}, \\ F(+\infty) - F(a), & \text{если } b = +\infty. \end{cases} \qquad (29.11)$$

Здесь, как всегда, $F(b-0) = \lim\limits_{x \to b-0} F(x)$ в случае, когда b конечно, и $F(+\infty) = \lim\limits_{x \to +\infty} F(x)$, а под первообразной F функции f на промежутке $[a, b)$ понимается, вообще говоря, обобщенная первообразная (см. п. 25.4*).

Равенство (29.11) понимается в том смысле, что либо обе его части одновременно имеют смысл, и тогда они равны, либо они одновременно не имеют смысла, т. е. стоящие в них пределы не существуют.

Согласно формуле Ньютона—Лейбница для функций, интегрируемых по Риману (см. теорему 5 п. 25.4*), для любого $\eta \in [a, b)$ имеем

$$\int_a^\eta f(x)\,dx = F(\eta) - F(a).$$

Перейдя в этом равенстве к пределу при $\eta \to b$, $a \leqslant \eta < b$, получим формулу (29.11).

Подчеркнем, что эта формула доказана в предположении, что функция f интегрируема в обычном смысле на каждом отрезке вида $[a, \eta)$, $a \leqslant \eta < b$. Для интегралов вида (29.8) в том случае, когда в правой части равенства более чем одно слагаемое, аналогичная формула верна не всегда. Образно говоря, если в некоторой внутренней точке данного промежутка функция обращается в бесконечность, то на всем этом промежутке нельзя, вообще говоря, применять, когда это возможно, формулу Ньютона—Лейбница. Например, если к интегралу $\int_{-1}^{1} \frac{dx}{x^2}$ формально применить формулу Ньютона—Лейбница, то он будет равен числу $-\frac{1}{x}\Big|_{-1}^{1} = -2$. Однако, как мы уже знаем, рассматриваемый интеграл не существует. Таким образом, в этом примере применение формулы Ньютона—Лейбница сразу на всем промежутке интегрирования невозможно по существу.

Формула, аналогичная (29.11), справедлива, конечно, для несобственных интегралов вида (29.6). Если же несобственный интеграл определяется равенством (29.8), то формулу Ньютона—Лейбница следует применять, когда это возможно, отдельно к каждому слагаемому правой части.

2^0 (линейность несобственного интеграла). *Если несобственные интегралы* $\int_a^b f(x)\,dx$, $\int_a^b g(x)\,dx$ *сходятся, то для любых чисел* λ, μ *сходится и несобственный интеграл* $\int_a^b [\lambda f(x) + \mu g(x)]\,dx$, *причем*

$$\int_a^b [\lambda f(x) + \mu g(x)]\,dx = \lambda\int_a^b f(x)\,dx + \mu\int_a^b g(x)\,dx.$$

В самом деле,

$$\int_a^b [\lambda f(x) + \mu g(x)]\,dx = \lim_{\eta \to b} \int_a^\eta [\lambda f(x) + \mu g(x)]\,dx =$$

$$= \lim_{\eta \to b} [\lambda \int_a^\eta f(x)\,dx + \mu \int_a^\eta g(x)\,dx] =$$

$$= \lambda \lim_{\eta \to b} \int_a^\eta f(x)\,dx + \mu \lim_{\eta \to b} \int_a^\eta g(x)\,dx =$$

$$= \lambda \int_a^b f(x)\,dx + \mu \int_a^b g(x)\,dx,\ a \leqslant \eta < b.$$

Аналогично доказываются и следующие ниже свойства несобственных интегралов, аналогичные соответствующим свойствам интеграла Римана.

Доказательства их предоставляются читателю.

3^0 (интегрирование неравенств). *Если интегралы* $\int_a^b f(x)\,dx$, $\int_a^b g(x)\,dx$ *сходятся и для всех* $x \in [a, b)$ *выполняется неравенство* $f(x) \leqslant g(x)$, *то*

$$\int_a^b f(x)\,dx \leqslant \int_a^b g(x)\,dx.$$

4^0 (правило интегрирования по частям). *Если функции* $u = u(x)$ *и* $v = v(x)$ *кусочно непрерывно дифференцируемы на каждом отрезке* $[a, \eta]$, $a < \eta < b$, *то*

$$\int_a^b u\,dv = uv\Big|_a^b - \int_a^b v\,du, \qquad (29.12)$$

причем если любые два из выражений $\int_a^b u\,dv$, $uv\Big|_a^b$ *и* $\int_a^b v\,du$ *имеют смысл (т. е. соответствующие пределы существуют и конечны), то имеет смысл и третье.*

5^0 (замена переменного в несобственном интеграле). *Пусть функция* $f(x)$ *непрерывна на полуинтервале* $[a, b)$, *функция* $\varphi(t)$ *непрерывно дифференцируема на полуинтервале* $[\alpha, \beta)$, $-\infty < \alpha < \beta \leqslant +\infty$, *причем* $a = \varphi(a) \leqslant \varphi(t) < b = \lim_{t \to \beta} \varphi(t)$ *при* $\alpha \leqslant t < \beta$; *тогда*

$$\int_a^b f(x)\,dx = \int_a^b f[\varphi(t)]\varphi'(t)\,dx. \qquad (29.13)$$

При этом интегралы в обеих частях этой формулы одновременно сходятся или нет.

Может случиться, что с помощью замены переменного несобственный интеграл превратится в обычный. Например, выполняя в несобственном интеграле $\int\limits_0^1 \frac{dx}{\sqrt{1 - x^2}}$ замену переменной $x = \sin t$, $0 \leqslant t < \pi/2$, получаем собственный интеграл

$$\int\limits_0^1 \frac{dx}{\sqrt{1 - x^2}} = \int\limits_0^{\pi/2} dt = \frac{\pi}{2}.$$

Отметим, что всякий несобственный интеграл $\int\limits_a^b f(x)\, dx$ по конечному промежутку $[a, b)$ может быть заменой переменной сведен к несобственному интегралу по неограниченному промежутку. Действительно, сделав, например, замену переменной

$$x = \frac{bt + a}{t + 1}, \quad dx = \frac{b - a}{(t + 1)^2}\, dt,$$

получим

$$\int\limits_a^b f(x)\, dx = (b - a) \int\limits_0^\infty f\left(\frac{bt + a}{t + 1}\right) \frac{dt}{(t + 1)^2}.$$

По аналогии с интегралом Римана для сходящегося несобственного интеграла $\int\limits_a^b f(x)\, dx$, $a < b$, по определению полагают

$$\int\limits_b^a f(x)\, dx = -\int\limits_a^b f(x)\, dx.$$

Следует обратить внимание на то, что не все свойства определенного интеграла Римана переносятся на несобственные интегралы. Так, например, произведение двух функций, интегрируемых по Риману на некотором отрезке, является функцией, также интегрируемой по Риману на нем. Аналог этого утверждения для несобственных интегралов несправедлив. Существуют функции f и g, интегралы от которых на некотором промежутке сходятся, а интеграл от их произведения на том же промежутке расходится. В самом деле, пусть, например, $f(x) = g(x) = 1/\sqrt{x}$. Как известно (см. п. 29.1), интеграл $\int\limits_0^1 \frac{dx}{\sqrt{x}}$ сходится, а интеграл

$$\int\limits_0^1 f(x)\, g(x)\, dx = \int\limits_0^1 \frac{dx}{x} \ \text{— расходится.}$$

Сделанное замечание еще раз напоминает о том, что, используя при обращении с несобственным интегралом аналогии свойств интеграла Римана, следует всегда не забывать о необходимости проверки справедливости для несобственного интеграла всякого утверждения, аналогичного соответствующему утверждению для собственного интеграла.

Примеры. Вычислим следующие несобственные интегралы, используя сформулированные выше свойства:

1. $\int\limits_1^{+\infty} \frac{dx}{x\sqrt{x^2-1}}$. С помощью замены переменной $x = \frac{1}{t}$ получим

$$\int\limits_1^{+\infty} \frac{dx}{x\sqrt{x^2-1}} = \int\limits_0^1 \frac{dt}{\sqrt{1-t^2}} = \arcsin t\Big|_0^1 = \frac{\pi}{2}.$$

2. $I_n = \int\limits_0^1 (\ln x)^n\,dx$, $n = 0, 1, 2, \ldots$. Интегрируя по частям (при $n > 0$), имеем

$$I_n = x(\ln x)^n\Big|_0^1 - n\int\limits_0^1 (\ln x)^{n-1}\,dx = -nI_{n-1},$$

так как $\lim\limits_{x\to+\infty} x(\ln x)^n = 0$. Это равенство легко получить, если применить n раз правило Лопиталя:

$$\lim_{x\to 0} x(\ln x)^n = \lim_{x\to 0}\frac{(\ln x)^n}{1/x} = -n\lim_{x\to 0}\frac{(\ln x)^{n-1}}{1/x} = \ldots =$$
$$= (-1)^{n+1}n!\lim_{x\to 0} x = 0.$$

Заметив, что $I_0 = \int\limits_0^1 dx = 1$, получим $I_n = (-1)^n n!$[1].

3. $J = \int\limits_0^{\pi/2} \ln\sin x\,dx$ (интеграл Эйлера).

Сделав замену переменного $x = 2t$, имеем

$$J = 2\int\limits_0^{\pi/4} \ln\sin 2t\,dt = 2\int\limits_0^{\pi/4} \ln(2\sin t\cos t)\,dt =$$
$$= \frac{\pi}{2}\ln 2 + 2\int\limits_0^{\pi/4} \ln\sin t\,dt + 2\int\limits_0^{\pi/4} \ln\cos t\,dt.$$

---

[1] Напомним, что, по определению, $0! = 1$.

Произведя в последнем интеграле замену переменного $t = \pi/2 - y$, получим

$$J = \frac{\pi}{2}\ln 2 + 2\int_0^{\pi/4} \ln\sin t\,dt + 2\int_{\pi/4}^{\pi/2} \ln\sin y\,dy = $$
$$= \frac{\pi}{2}\ln 2 + 2\int_0^{\pi/2} \ln\sin t\,dt,$$

т. е. $J = \frac{\pi}{2}\ln 2 + 2J$, откуда $J = -\frac{\pi}{2}\ln 2$.

4. $J_n = \int_0^{+\infty} x^n e^{-x}\,dx$, $n = 0, 1, 2, \ldots$. Снова проинтегрировав по частям заданный интеграл при $n > 0$, получим

$$J_n = -x^n e^{-x}\Big|_0^{+\infty} + n\int_0^{+\infty} x^{n-1}e^{-x}\,dx = nJ_{n-1},$$

и так как

$$J_0 = \int_0^{+\infty} e^{-x}\,dx = -e^{-x}\Big|_0^{+\infty} = 1,$$

то $J_n = n!$.

5. Остаются справедливыми для несобственных интегралов неравенства Минковского и Гёльдера (см. п. 28.2*):

$$\left(\int_a^b |f(x) + g(x)|^p\,dx\right)^{\frac{1}{p}} \leqslant \left(\int_a^b |f(x)|^p\,dx\right)^{\frac{1}{p}} + \left(\int_a^b |g(x)|^p\,dx\right)^{\frac{1}{p}},$$

$$\left|\int_a^b f(x)g(x)\,dx\right| \leqslant \left(\int_a^b |f(x)|^p\,dx\right)^{\frac{1}{p}} \left(\int_a^b |g(x)|^q\,dx\right)^{\frac{1}{p}},$$

$$1 < p < +\infty,\ \frac{1}{p} + \frac{1}{q} = 1.$$

Для доказательства достаточно написать соответствующие неравенства для интегралов на отрезке $[a, \eta]$ и перейти к пределу при $\eta \to b$.

В следующем пункте будет рассмотрена специфическая задача теории несобственных интегралов, а именно: приведен ряд условий, при выполнении которых несобственные интегралы заведомо сходятся (расходятся).

29.3. Несобственные интегралы от неотрицательных функций

Изучение признаков сходимости несобственных интегралов начнем со случая, когда подынтегральная функция неотрицательна. При этом будем придерживаться соглашения, сформулированного в начале предыдущего пункта.

ЛЕММА. *Если функция f неотрицательна на полуинтервале $[a, b)$, то для сходимости несобственного интеграла $\int_a^b f(x)\,dx$ необходимо и достаточно, чтобы все интегралы $\int_a^\eta f(x)\,dx$, $a \leqslant \eta < b$, были ограниченными в совокупности, т. е. чтобы существовала такая постоянная $M > 0$, что для всех $\eta \in [a, b)$ выполняется неравенство*

$$\int_a^\eta f(x)\,dx \leqslant M. \qquad (29.14)$$

При выполнении этого условия

$$\int_a^b f(x)\,dx = \sup_{a \leqslant \eta < b} \int_a^\eta f(x)\,dx. \qquad (29.15)$$

Доказательство. Рассмотрим функцию

$$\varphi(\eta) = \int_a^\eta f(x)\,dx, \quad a \leqslant \eta < b. \qquad (29.16)$$

В силу того что $f \geqslant 0$, функция φ возрастает: действительно, если $a \leqslant \eta < \eta' < b$, то (см. свойство 8^0 интеграла в п. 24.1)

$$\int_\eta^{\eta'} f(x)\,dx \geqslant 0,$$

поэтому

$$\varphi(\eta') = \int_a^{\eta'} f(x)\,dx = \int_a^\eta f(x)\,dx + \int_\eta^{\eta'} f(x)\,dx \geqslant \int_a^\eta f(x)\,dx = \varphi(\eta).$$

Теперь заметим, что несобственный интеграл $\int_a^b f(x)\,dx$ сходится тогда и только тогда, когда существует конечный предел $\lim\limits_{\eta \to b} \int_a^\eta f(x)\,dx = \lim\limits_{\eta \to b} \varphi(\eta)$, а последний существует в том и только том случае (см. теорему 5 в п. 5.14), когда функция φ ограничена сверху, т. е. когда выполняется условие (29.14). При этом

$$\int_a^b f(x)\,dx = \lim_{\eta \to b} \varphi(\eta) = \sup_{a \leqslant \eta < b} \varphi(\eta) = \sup_{a \leqslant \eta \leqslant b} \int_a^\eta f(x)\,dx. \ \square$$

Из доказанной леммы следует, что, для того чтобы несобственный интеграл $\int_a^b f(x)\,dx$ от неотрицательной функции расходился, необходимо и достаточно, чтобы функция $\varphi(\eta)$ (см. (29.16)) была неограниченной сверху; но тогда, в силу ее возрастания,

$$\lim_{\eta \to b} \int_a^\eta f(x)\,dx = \lim_{\eta \to b} \varphi(\eta) = +\infty.$$

Поэтому если несобственный интеграл $\int_a^b f(x)\,dx$ от неотрицательной функции расходится, то пишут $\int_a^b f(x)\,dx = +\infty$. При таком соглашении остается справедливым равенство (29.15).

ТЕОРЕМА 1 (признак сравнения). *Пусть функции f и g неотрицательны на полуинтервале $[a, b)$ и*

$$f(x) = O(g(x)) \text{ при } x \to b^{1}. \tag{29.17}$$

Тогда:

1) *если интеграл $\int_a^b g(x)\,dx$ сходится, то сходится и интеграл $\int_a^b f(x)\,dx$;*

2) *если интеграл $\int_a^b f(x)\,dx$ расходится, то расходится и интеграл $\int_a^b g(x)\,dx$.*

СЛЕДСТВИЕ. *Пусть функции f, g неотрицательны на полуинтервале $[a, b)$, $g(x) \neq 0$, $x \in [a, b)$ и существует*

$$\lim_{x \to b} \frac{f(x)}{g(x)} = k,\ a \leqslant x < b. \tag{29.18}$$

Тогда:

1) *если интеграл $\int_a^b g(x)\,dx$ сходится и $0 \leqslant k < +\infty$, то интеграл $\int_a^b f(x)\,dx$ также сходится;*

2) *если интеграл $\int_a^b g(x)\,dx$ расходится и $0 < k \leqslant +\infty$, то интеграл $\int_a^b f(x)\,dx$ также расходится.*

В частности, если f и g — эквивалентные при $x \to b$ функции: $f \sim g$, $x \to b$ (см. п. 8.2), *то интегралы $\int_a^b f(x)\,dx$ и $\int_a^b g(x)\,dx$ сходятся или расходятся одновременно.*

Доказательство теоремы. Пусть интеграл $\int_a^b g(x)\,dx$ сходится. Из условия (29.17) следует существование такого η_0,

[1] В частности, $f(x) \leqslant g(x)$, $x \in [a, b)$.

$a \leqslant \eta_0 < b$, и такого $c > 0$, что для всех $x \in [\eta_0, b)$ выполняется неравенство

$$f(x) \leqslant cg(x) \tag{29.19}$$

(см. п. 8.2). Из сходимости интеграла $\int\limits_a^b g(x)\,dx$ следует и сходимость интеграла $\int\limits_{\eta_0}^{\eta} g(x)\,dx$. В силу же необходимости условий леммы для сходимости интеграла, существует такое число $M > 0$, что для любого $\eta \in [\eta_0, b)$ справедливо неравенство $\int\limits_{\eta_0}^{\eta} g(x)dx \leqslant M$. Отсюда и из неравенства (29.19) имеем

$$\int\limits_{\eta_0}^{\eta} f(x)\,dx \leqslant c\int\limits_{\eta_0}^{\eta} g(x)\,dx \leqslant cM.$$

Из этого неравенства, в силу достаточности условий леммы для сходимости интеграла от неотрицательной функции, получаем, что интеграл $\int\limits_{\eta_0}^{b} f(x)\,dx$, следовательно, и интеграл $\int\limits_a^b f(x)\,dx$ сходятся.

Первое утверждение теоремы доказано. Второе — логически равносильно первому: если интеграл $\int\limits_a^b f(x)\,dx$ расходится, то $\int\limits_a^b g(x)\,dx$ не может сходиться, так как если бы он был сходящимся, то, в силу уже доказанного первого утверждения теоремы, сходился бы и интеграл $\int\limits_a^b f(x)\,dx$. Таким образом, интеграл $\int\limits_a^b g(x)\,dx$ расходится. □

Доказательство следствия. Из выполнения условия (29.18) для k, удовлетворяющего условию $0 \leqslant k < +\infty$, следует, что существует такое $\eta \in [a, b)$, что если $\eta < x < b$, то

$$\frac{f(x)}{g(x)} < k + 1, \text{ т. е. } f(x) < (k+1)g(x),$$

а это означает, что

$$f(x) = O(g(x)),\ x \to b.$$

Поэтому утверждение 1) следствия непосредственно вытекает из утверждения 1) теоремы 1.

Пусть теперь условие (29.18) выполнено при некотором k, удовлетворяющем условию $0 < k \leqslant +\infty$. Тогда для любого $k' \in (0, k)$ существует такое $\eta \in [a, b)$, что если $\eta < x < b$, то

$$\frac{f(x)}{g(x)} > k', \text{ или } g(x) < \frac{1}{k'} f(x).$$

Это и означает, что $g(x) = O(f(x))$, $x \to b$. Поэтому утверждение 2) следствия непосредственно вытекает из утверждения 2) теоремы 1. □

Функция $g(x)$ в утверждении 1) теоремы 1 и в ее следствии, с помощью которой устанавливается сходимость интеграла $\int\limits_a^b f(x)\,dx$, называется *функцией сравнения*. Если, в частности, $f(x) \leqslant g(x)$ для всех $x \in [a, b)$, то говорят также, что $f(x)$ мажорируется функцией $g(x)$ или что $g(x)$ служит *мажорантой* для $f(x)$.

Эффективность использования критерия сравнения для решения вопроса о сходимости интеграла зависит, конечно, от запаса функций сравнения, о которых известно, сходится или расходится несобственный интеграл от них, взятый по рассматриваемому промежутку, и которые, тем самым, можно пытаться использовать для исследования сходимости данного интеграла. Заметим, что утверждение, аналогичное теореме 1, справедливо, конечно, и для несобственных интегралов типа (29.6).

В качестве функций сравнения $g(x)$ часто достаточно брать степенные функции. Именно, в случае конечных промежутков $[a, b)$ и $(a, b]$ берутся, соответственно, функции $g(x) = \dfrac{1}{(b-x)^\alpha}$ и $g(x) = \dfrac{1}{(x-a)^\alpha}$, интегралы от которых $\int\limits_a^b \dfrac{dx}{(b-x)^\alpha}$, $\int\limits_a^b \dfrac{dx}{(x-a)^\alpha}$ сходятся при $\alpha < 1$ и расходятся при $\alpha \geqslant 1$ (в этом легко убедиться, сведя указанные интегралы линейной заменой переменной к интегралам $\int\limits_0^1 \dfrac{dx}{x^\alpha}$, рассмотренным в п. 29.1). В случае же бесконечных промежутков $[a, +\infty)$ и $(-\infty, b]$ за функции сравнения берутся, соответственно, $g(x) = \dfrac{1}{x^\alpha}$ и $g(x) = \dfrac{1}{|x|^\alpha}$, интегралы от которых $\int\limits_1^{+\infty} \dfrac{dx}{x^\alpha}$ и $\int\limits_{+\infty}^{-1} \dfrac{dx}{|x|^\alpha}$ сходятся при $\alpha > 1$ и расходятся при $\alpha \leqslant 1$ (см. примеры в п. 29.1).

Отметим еще, что все сформулированные признаки сходимости и расходимости интегралов остаются в силе (с очевидными изменениями), если в них условие неотрицательности функции f заменить условием ее неположительности (это следует из того, что интеграл $\int_a^b (-f(x))\,dx$ сходится тогда и только тогда, когда сходится интеграл $\int_a^b f(x)\,dx$).

Примеры. 1. Интеграл

$$\int_0^1 \frac{x^2}{\sqrt[3]{1-x^2}}\,dx \tag{29.20}$$

сходится. В самом деле, обозначая через f подынтегральную функцию: $f(x) = \dfrac{x^2}{\sqrt[3]{1-x^2}}$ и используя функцию сравнения $g(x) = \dfrac{1}{\sqrt[3]{1-x}}$, $\alpha = \dfrac{1}{3}$, имеем

$$\lim_{x \to 1-0} \frac{f(x)}{g(x)} = \lim_{x \to 1-0} \sqrt[3]{1-x}\,\frac{x^2}{\sqrt[3]{1-x^2}} = \lim_{x \to 1-0} \frac{x^2}{\sqrt[3]{1+x}} = \frac{1}{\sqrt[3]{2}},$$

поэтому, согласно следствию из теоремы 1, интеграл (29.20) сходится.

2. Интеграл $\int_0^1 \frac{1}{1-x^2}\,dx$ расходится. Чтобы убедиться в этом, достаточно взять в качестве функции сравнения $g(x) = \dfrac{1}{1-x}$, $\alpha = 1$.

В рассмотренных примерах показатель α у функции сравнения можно было выбрать сразу, исходя из конкретного вида заданной подынтегральной функции. Иногда, когда такой выбор сразу не ясен, приходится предварительно проделывать некоторые дополнительные исследования, например, попытаться выделить ее главную часть, прибегнув к формуле Тейлора. Рассмотрим подобные примеры.

3. Интеграл

$$\int_0^1 \ln x\,dx \tag{29.21}$$

сходится. Действительно, по правилу Лопиталя при любом $\alpha > 0$, в частности при $0 < \alpha < 1$, имеем

$$\lim_{x \to +0} \frac{\ln x}{1/x^{\alpha}} = \lim_{x \to +0} \frac{1/x}{\alpha x^{-\alpha-1}} = -\frac{1}{\alpha} \lim_{x \to +0} x^{\alpha} = 0,$$

поэтому, согласно следствию из теоремы 1 (точнее, его аналогу для неположительных функций), интеграл (29.21) сходится.

Это было установлено и ранее непосредственным вычислением данного интеграла (см. пример 2 в п. 29.2).

Геометрически сходимость и расходимость интегралов (29.9), (29.10) и (29.21) означает конечность или бесконечность площадей соответствующих «бесконечных криволинейных трапеций», сравнительное расположение которых изображено на рисунке 142.

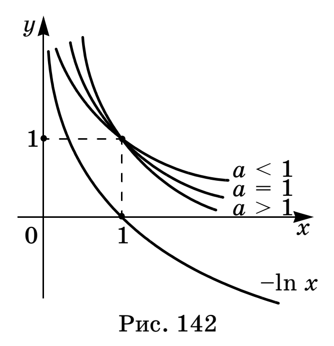


Рис. 142

4. Для выяснения вопроса о сходимости интеграла

$$\int_0^1 \frac{dx}{\ln x} \qquad (29.22)$$

заметим, что $\ln x = \ln [1 + (x - 1)] \sim x - 1$ при $x \to 1$, и возьмем за функцию сравнения $g(x) = \dfrac{1}{x-1}$, т. е. выберем $\alpha = 1$. Тогда $\lim\limits_{x \to 1} \dfrac{x-1}{\ln x} = 1$ и, следовательно, интеграл (29.22) расходится.

5. Интеграл

$$\int_1^{+\infty} \frac{\ln x}{\sqrt{x^3+1}}\, dx \qquad (29.23)$$

сходится. Действительно, возьмем $\alpha = 3/2 - \varepsilon$, $\varepsilon > 0$. Тогда, применив снова правило Лопиталя, получим

$$\lim_{x \to +\infty} \frac{x^{3/2-\varepsilon}\ln x}{\sqrt{x^3+1}} = \lim_{x \to +\infty} \frac{x^{3/2}}{\sqrt{x^3+1}} \lim_{x \to +\infty} \frac{\ln x}{x^{\varepsilon}} = \lim_{x \to +\infty} \frac{1}{\varepsilon x^{\varepsilon}} = 0.$$

Выберем $\varepsilon > 0$ так, чтобы $3/2 - \varepsilon > 1$; в этом случае интеграл $\int_1^{+\infty} \dfrac{dx}{x^{\frac{3}{2}-\varepsilon}}$ сходится, а поэтому, в силу следствия из теоремы 1, сходится и интеграл (29.23).

6. Исследуем сходимость интеграла

$$\int_1^{+\infty} \frac{\ln\cos\frac{1}{x}}{x^p}\, dx. \qquad (29.24)$$

Здесь подынтегральная функция всюду отрицательна. Очевидно, интеграл (29.24) сходится или расходится одновременно с интегралом

$$\int_1^{+\infty}\left(-\frac{\ln\cos\frac{1}{x}}{x^p}\right)dx, \qquad (29.25)$$

у которого подынтегральная функция всюду положительна. Разложив функцию $\ln\cos\frac{1}{x}$ по формуле Тейлора, получим

$$-\frac{1}{x^p}\ln\cos\frac{1}{x} = -\frac{\ln\left[1-\frac{1}{2x^2}+o\left(\frac{1}{x^2}\right)\right]}{x^p} = -\frac{-\frac{1}{2x^2}+o\left(\frac{1}{x^2}\right)}{x^p} =$$

$$= \frac{1}{2x^{2+p}} + o\left(\frac{1}{x^{2+p}}\right), \quad x \to +\infty.$$

Таким образом, $-\frac{\ln\cos(1/x)}{x^p} \sim \frac{1}{2x^{2+p}}$ при $x \to +\infty$ и, следовательно, интеграл (29.24) сходится при $2 + p > 1$, т. е. при $p > -1$, и расходится при $p \leqslant -1$.

В примерах 2 и 3 сходимость рассмотренных там интегралов можно было бы установить, вычислив их по формуле Ньютона—Лейбница. Однако выяснение сходимости интегралов с помощью признака сравнения обычно требует меньше вычислений, чем при их предварительном нахождении по формуле Ньютона—Лейбница. Важно отметить, что, используя признак сравнения, можно выяснить сходимость интегралов, конечно, и в случае, когда первообразная подынтегральной функции не является элементарной, и, следовательно, обычным приемом (с помощью формулы Ньютона—Лейбница) интеграл заведомо не вычисляется, как это и было в примерах 4 и 5.

УПРАЖНЕНИЯ.

Исследовать на сходимость следующие интегралы.

4. $\int_0^{+\infty}\frac{x^2dx}{x^4+x^2+1}$.

5. $\int_0^1\frac{dx}{\sqrt{1-x^2}}$.

6. $\int_{-\infty}^{+\infty}\frac{dx}{\sqrt{1+x^4}}$.

7. $\int_0^1\frac{dx}{x+\sqrt{x}}$.

8. $\int\limits_0^{+\infty} x^3 e^{-x}\,dx.$

9. $\int\limits_1^{+\infty} \frac{e^x}{x^{10}}\,dx.$

10. $\int\limits_0^1 \frac{\ln x}{\sqrt[3]{x}}\,dx.$

11. $\int\limits_1^{+\infty} \frac{dx}{x\ln x}.$

12. $\int\limits_1^{+\infty} \frac{dx}{x\ln^2 x}.$

13. $\int\limits_2^{+\infty} \frac{dx}{x\ln^2 x}.$

14. $\int\limits_0^{\pi/2} \operatorname{tg}^{\alpha} x\,dx,\ -\infty < \alpha < +\infty.$

15. $\int\limits_1^{+\infty} \frac{dx}{1 + (\ln x)^p},\ -\infty < p < +\infty.$

16. $\int\limits_0^{\pi} \frac{\ln\sin x}{\sqrt{x}}\,dx$

Подчеркнем еще раз, что признак сравнения для выяснения вопроса о сходимости несобственного интеграла можно применять только для функций, не меняющих знака. Возникает вопрос: как выяснить, сходится или расходится несобственный интеграл в случае, когда подынтегральная функция меняет знак? В следующих пунктах этот вопрос и будет изучен.

29.4. Критерий Коши сходимости несобственных интегралов

В этом пункте мы уже не будем предполагать, что значения рассматриваемых функций сохраняют один и тот же знак на полуинтервале $[a, b)$, — они могут принимать значения любого знака, но по-прежнему будем предполагать, что все рассматриваемые функции при любом выборе числа $\eta \in [a, b)$ интегрируемы по Риману на отрезке $[a, \eta]$.

ТЕОРЕМА 2. *Для сходимости интеграла* $\int\limits_a^b f(x)\,dx$ *необходимо и достаточно, чтобы для любого* $\varepsilon > 0$ *существовало такое число* η, $a \leqslant \eta < b$, *что если* $\eta < \eta' < b$, $\eta < \eta'' < b$, *то*

$$\left|\int\limits_{\eta''}^{\eta'} f(x)\,dx\right| < \varepsilon. \tag{29.26}$$

Доказательство. Положим

$$\varphi(\eta) = \int\limits_a^{\eta} f(x)\,dx,\ \ a \leqslant \eta < b \leqslant +\infty.$$

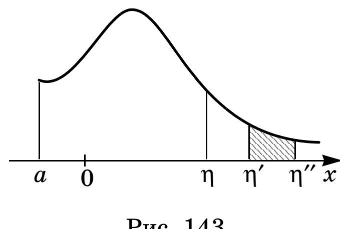


Рис. 143

Тогда сходимость интеграла $\int\limits_a^b f(x)\,dx$, т. е. существование конечного предела (29.1), означает существование конечного предела $\lim\limits_{\eta \to b} \varphi(\eta)$. В силу же критерия Коши для наличия конечного предела функции $\varphi(\eta)$ при $\eta \to b$ необходимо и достаточно, чтобы для любого $\varepsilon > 0$ существовала такая левосторонняя проколотая окрестность

$$\overset{\circ}{U}(b) = \{x : \eta < x < b\}$$

точки b, т. е. существовало такое число η, $a \leqslant \eta < b$, что для всех $\eta' \in \overset{\circ}{U}(b)$ и $\eta'' \in \overset{\circ}{U}(b)$ (что равносильно условию $\eta < \eta' < b$, $\eta < \eta'' < b$) выполнялось бы неравенство

$$|\varphi(\eta'') - \varphi(\eta')| < \varepsilon. \qquad (29.27)$$

Так как

$$\varphi(\eta'') - \varphi(\eta') = \int\limits_a^{\eta'} f(x)\,dx - \int\limits_a^{\eta''} f(x)\,dx = \int\limits_{\eta'}^{\eta''} f(x)\,dx,$$

то неравенство (29.27) равносильно условию (29.26) (рис. 143). □

Теорема 2 называется *критерием Коши сходимости интеграла*.

29.5. Абсолютно сходящиеся интегралы

Важным понятием для несобственных интегралов от функций, меняющих знак, является понятие абсолютно сходящегося интеграла.

Определение 2. *Несобственный интеграл* $\int\limits_a^b f(x)\,dx$ *называется абсолютно сходящимся, если сходится интеграл* $\int\limits_a^b |f(x)|\,dx$.

Функции, для которых интеграл $\int\limits_a^b f(x)\,dx$ абсолютно сходится, называются абсолютно интегрируемыми (в несобственном смысле) на промежутке с концами a и b. В случае,

когда a и b конечны, говорят также, что функция f абсолютно интегрируема на отрезке $[a, b]$.

Из теоремы 2 непосредственно следует критерий абсолютной сходимости интеграла.

ТЕОРЕМА 3. *Для того чтобы интеграл* $\int_a^b f(x)\,dx$ *абсолютно сходился, необходимо и достаточно, чтобы для любого* $\varepsilon > 0$ *существовало такое* $\eta = \eta(\varepsilon)$, *что если* $\eta < \eta' < b$ *и* $\eta < \eta'' < b$, *то*

$$\left|\int_{\eta'}^{\eta''} |f(x)\,dx\right| < \varepsilon.$$

Эта теорема называется *критерием Коши абсолютной сходимости интеграла*.

Напомним, что, как всегда, здесь предполагается, что функция f интегрируема по Риману на любом отрезке $[a, \eta]$, где $a \leqslant \eta < b$, $-\infty < a < b \leqslant +\infty$.

Признак сходимости интегралов от неотрицательных функций, очевидно, применим также и для выяснения абсолютной сходимости интегралов. Пусть, например, требуется выяснить: сходится абсолютно или нет интеграл

$$\int_1^{+\infty} \frac{\cos x}{x^2}\,dx. \tag{29.28}$$

Из неравенства $\left|\frac{\cos x}{x^2}\right| \leqslant \frac{1}{x^2}$ и сходимости интеграла $\int_1^{+\infty} \frac{dx}{x^2}$, согласно признаку сравнения, следует и сходимость интеграла $\int_1^{+\infty} \left|\frac{\cos x}{x^2}\right| dx$. Это означает, что интеграл (29.28) абсолютно сходится.

Важную связь между сходимостью и абсолютной сходимостью интегралов устанавливает следующая теорема.

ТЕОРЕМА 4. *Если интеграл абсолютно сходится, то он и просто сходится.*

Доказательство. Пусть задано $\varepsilon > 0$. Если интеграл $\int_a^b f(x)\,dx$ абсолютно сходится, то, в силу критерия Коши абсолютной сходимости интеграла (см. теорему 3), для любого $\varepsilon > 0$ существует такое η, $a < \eta < b$, что если $\eta < \eta' < b$, $\eta < \eta'' < b$, то

$$\left|\int_{\eta'}^{\eta''} |f(x)|\,dx\right| < \varepsilon. \tag{29.29}$$

Так как $\left|\int\limits_{\eta'}^{\eta''} f(x)\,dx\right| \leqslant \left|\int\limits_{\eta'}^{\eta''} |f(x)|\,dx\right|$, то, в силу неравенства (29.29), для любых указанных η' и η'' имеем $\left|\int\limits_{\eta'}^{\eta''} f(x)\,dx\right| < \varepsilon$, поэтому, в силу критерия Коши сходимости интегралов (см. теорему 2), интеграл $\int\limits_a^b f(x)\,dx$ сходится. □

УПРАЖНЕНИЕ 17. Если несобственный интеграл от функции, определенной на отрезке, абсолютно сходится, то он и просто сходится. Интеграл Римана является частным случаем несобственного интеграла. Следовательно, если существует интеграл Римана от абсолютной величины функции, то существует и интеграл Римана от самой функции. Это неверно (привести соответствующий пример!). Где ошибка в проведенном рассуждении?

Существенно отметить, что интеграл может сходиться, но не сходиться абсолютно (в этом случае его иногда называют *условно сходящимся интегралом*).

В качестве примера рассмотрим интеграл

$$\int\limits_0^{+\infty} \frac{\sin x}{x}\,dx. \tag{29.30}$$

Прежде всего заметим, что $\lim\limits_{x \to 0} \frac{\sin x}{x} = 1$, поэтому подынтегральная функция, доопределенная единицей при $x = 0$, непрерывна на полупрямой $x \geqslant 0$ и, значит, интегрируема по Риману на любом отрезке $[0, \eta]$, в частности на отрезке $[0, 1]$. Поэтому вопрос о сходимости (абсолютной сходимости) интеграла (29.30) эквивалентен вопросу о сходимости (абсолютной сходимости) интеграла

$$\int\limits_1^{+\infty} \frac{\sin x}{x}\,dx. \tag{29.31}$$

Для исследования его сходимости выполним интегрирование по частям:

$$\int\limits_1^{+\infty} \frac{\sin x}{x}\,dx = -\int\limits_1^{+\infty} \frac{1}{x}\,d(\cos x) =$$

$$= \frac{\cos x}{x}\Bigg|_1^{+\infty} + \int\limits_1^{+\infty} \cos x\,d\left(\frac{1}{x}\right) = \cos 1 - \int\limits_1^{+\infty} \frac{\cos x}{x^2}\,dx.$$

В правой части равенства получен интеграл (29.28), который, как известно, абсолютно, а значит, и просто сходится.

Таким образом, оба выражения в правой части равенства имеют смысл и, следовательно, конечны. Поэтому, во-первых, проделанное интегрирование по частям законно, а во-вторых, левая часть равенства также конечна, т. е. интеграл (29.31) сходится.

Заметим, что в результате интегрирования по частям мы заменили интеграл (29.31) суммой некоторого конечного выражения и другого несобственного интеграла, у которого в знаменателе подынтегральной функции стоит более высокая степень переменной интегрирования, чем в интеграле (29.31), а в числителе — ограниченная, как и в (29.31), функция. В получившемся интеграле подынтегральная функция быстрее стремится к нулю, чем в исходном, в том смысле, что $\frac{1}{x^2} = o\left(\frac{1}{x}\right)$ при $x \to \infty$. Поэтому его сходимость оказалось легче непосредственно исследовать, чем сходимость исходного интеграла: он оказался даже не просто сходящимся, а абсолютно сходящимся.

Метод, позволяющий свести исследование сходимости данного интеграла к исследованию сходимости другого интеграла, который в каком-то смысле «лучше сходится», чем данный, называется *методом улучшения сходимости.*

Покажем теперь, что интеграл (29.31) не сходится абсолютно, т. е. что интеграл

$$\int_1^{+\infty} \frac{|\sin x|}{x}\,dx \qquad (29.32)$$

расходится. Действительно, из неравенства

$$|\sin x| \geqslant \sin^2 x = \frac{1 - \cos 2x}{2}$$

при любом $\eta > 1$ имеем

$$\int_1^{\eta} \frac{|\sin x|}{x}\,dx \geqslant \frac{1}{2}\int_1^{\eta} \frac{dx}{x} - \frac{1}{2}\int_1^{\eta} \frac{\cos 2x}{x}\,dx. \qquad (29.33)$$

Интеграл $\int_1^{+\infty} \frac{dx}{x}$ расходится и равен $+\infty$. Интеграл же $\int_1^{+\infty} \frac{\cos 2x}{x}\,dx$ сходится. Чтобы в этом убедиться, проинтегрируем его по частям:

$$\int_1^{+\infty} \frac{\cos 2x}{x}\,dx = \frac{1}{2}\int_1^{+\infty} \frac{1}{x}\,d(\sin 2x) =$$

$$= \frac{\sin 2x}{2x}\bigg|_1^{+\infty} - \frac{1}{2}\int_1^{+\infty} \sin 2x\,d\left(\frac{1}{x}\right) = -\frac{\sin 2}{2} + \frac{1}{2}\int_1^{+\infty} \frac{\sin 2x}{x^2}\,dx.$$

В силу этой формулы, сходимость интеграла $\int\limits_1^{+\infty} \frac{\cos 2x}{x}\, dx$ непосредственно следует из абсолютной сходимости интеграла $\int\limits_1^{+\infty} \frac{\sin 2x}{x^2}\, dx$, которая, в свою очередь, вытекает из очевидного неравенства $\left|\frac{\sin 2x}{x^2}\right| \leqslant \frac{1}{x^2}$. Перейдя теперь к пределу при $\eta \to +\infty$ в неравенстве (29.33), получаем, что правая, а следовательно, и левая части этого неравенства стремятся к $+\infty$ и поэтому интеграл (29.32) расходится.

Таким образом, интеграл (29.31), значит, и интеграл (29.30) не сходятся абсолютно.

Докажем еще одно полезное для дальнейшего вспомогательное утверждение.

ЛЕММА 2. *Если функция f абсолютно интегрируема, а функция g интегрируема по Риману на отрезке* $[a, b]$, *то их произведение gf также абсолютно интегрируемо на* $[a, b]$.

Доказательство. Как было оговорено выше, рассматриваются только такие функции f, которые при любом $\eta \in [a, b)$ интегрируемы по Риману на отрезке $[a, \eta]$. По условию, функция g интегрируема по Риману на отрезке $[a, b]$, следовательно, она интегрируема по Риману и на всяком отрезке $[a, \eta]$, $\eta \in [a, b)$ (см. свойство 2^0 в п. 24.1). Поэтому произведение gf также интегрируемо по Риману на любом указанном отрезке $[a, \eta]$ (см. свойство 6^0 в п. 24.1). Это означает, что имеет смысл рассмотрение несобственного интеграла $\int\limits_a^b g(x)f(x)\, dx$.

В силу интегрируемости по Риману функции g на отрезке $[a, b]$, она ограничена на нем, т. е. существует такая постоянная $M > 0$, что для всех $x \in [a, b]$ выполняется неравенство $|g(x)| \leqslant M$. Следовательно, для всех $x \in [a, b)$ справедливо и неравенство $|g(x)f(x)| \leqslant M|f(x)|$. Заметив, что, в силу абсолютной интегрируемости функции f на отрезке $[a, b]$, интеграл $\int\limits_a^b M|f(x)|\, dx = M\int\limits_a^b |f(x)|\, dx$ сходится, получим, согласно признаку сравнения, что сходится и интеграл $\int\limits_a^b |g(x)f(x)|\, dx$,

т. е. что произведение gf абсолютно интегрируемо на отрезке $[a, b]$. □

Все сказанное в этом пункте естественным образом переносится и на несобственные интегралы других видов, рассмотренных в п. 29.1, т. е. на интегралы вида (29.6), а также на интегралы общего типа (29.8).

29.6. Исследование сходимости интегралов

Докажем один достаточный признак сходимости интегралов, называемый обычно *признаком Дирихле*.

ТЕОРЕМА 5 (признак Дирихле). *Пусть*:

1) *функция f непрерывна и имеет ограниченную первообразную F при $x \geqslant a$*;

2) *функция g непрерывно дифференцируема и убывает при $x \geqslant a$*;

3) $\lim\limits_{x \to +\infty} g(x) = 0$.

Тогда сходится интеграл

$$\int_a^{+\infty} f(x)g(x)\,dx. \tag{29.34}$$

Доказательство. Прежде всего заметим, что, в силу сделанных предположений, функция fg непрерывна, а значит, и интегрируема по Риману на любом отрезке $[a, b]$, $a < b < +\infty$, и поэтому имеет смысл говорить о несобственном интеграле (29.34).

Проинтегрировав по частям произведение $f(x)g(x)$ на отрезке $[a, b]$, получим

$$\int_a^b f(x)g(x)\,dx = \int_a^b g(x)\,dF(x) = g(x)\,F(x)\Big|_a^b - \int_a^b F(x)g'(x)\,dx. \tag{29.35}$$

Исследуем поведение обоих слагаемых правой части при $b \to +\infty$. В силу ограниченности функции F (см. условие 1 теоремы),

$$M = \sup |F(x)| < +\infty.$$

Из условий 2 и 3 теоремы следует, что функция g не отрицательна для всех $x \geqslant a$, в частности $g(b) \geqslant 0$; поэтому

$$|g(b)F(b)| \leqslant Mg(b).$$

Кроме того, в силу условия 3,

$$\lim_{b \to +\infty} g(b)F(b) = 0,$$

и, следовательно,

$$\lim_{b \to +\infty} g(x)F(x)\Big|_a^b = -g(a)F(a).$$

Далее, из убывания функции g следует, что $g'(x) \leqslant 0$ при $x \geqslant a$, поэтому

$$\int_a^b |F(x)g'(x)|\, dx \leqslant M\int_a^b |g'(x)|\, dx =$$

$$= -M\int_a^b g'(x)\, dx = M[g(a) - g(b)] \leqslant Mg(a),$$

так как $g(b) \geqslant 0$.

Таким образом, интегралы $\int_a^b |F(x)g'(x)|\, dx$ ограничены в совокупности при всех $b > a$, поэтому интеграл $\int_a^{+\infty} F(x)g'(x)\, dx$ абсолютно, а значит, и просто сходится, т. е. существует конечный предел $\lim\limits_{b \to +\infty} \int_a^b F(x)g'(x)\, dx$.

Мы доказали, что в правой части равенства (29.35) оба слагаемых при $b \to +\infty$ имеют конечный предел, следовательно, предел левой части при $b \to \infty$ также конечен. Это и означает сходимость интеграла (29.34). □

З а м е ч а н и е 1. Получение нужных оценок в приведенном доказательстве напоминает рассуждения при доказательстве второй интегральной теоремы о среднем (см. п. 26.3*). Это не случайно: если воспользоваться указанной теоремой для оценки интеграла $\int_a^b f(x)g(x)\, dx$, то признак Дирихле можно доказать короче. Мы не стали этого делать, чтобы еще раз показать, как с помощью интегрирования по частям можно улучшить сходимость интеграла.

ТЕОРЕМА 6 (признак Абеля[1]). *Если на полуоси $x \geqslant a$:*

1) *функция f непрерывна и сходится интеграл*

$$\int_a^{+\infty} f(x)\, dx; \tag{29.36}$$

2) *функция g непрерывно дифференцируема, ограничена и монотонна, то интеграл $\int_a^{+\infty} f(x)g(x)\, dx$ сходится.*

[1] Н. Абель (1802—1829) — норвежский математик.

Доказательство. Покажем, что эта теорема вытекает из предыдущей. Прежде всего отметим, что интегралы

$$\int_a^{+\infty} f(x)g(x)\,dx \text{ и } \int_a^{+\infty} f(x)[-g(x)]\,dx$$

сходятся или расходятся одновременно и что, в силу монотонности функции g, одна из функций g или $-g$ убывает.

Пусть, для определенности, убывает функция g. В силу ее ограниченности и монотонности, существует конечный предел $\lim\limits_{x \to +\infty} g(x) = c$, а так как функция g убывает, то при $x \to +\infty$, убывая, стремится к нулю и разность $g(x) - c$.

Представим произведение $f(x)g(x)$ в виде

$$f(x)g(x) = f(x)[g(x) - c] + cf(x). \qquad (29.37)$$

В силу сходимости интеграла (29.36), интеграл $\int_a^{+\infty} cf(t)\,dx$ также сходится. Из этого же условия следует, что интегралы $F(x) = \int_a^x f(t)\,dt$, $x \geqslant a$, ограничены. В самом деле, из существования конечного предела

$$\lim_{x \to +\infty} F(x) = \int_a^{+\infty} f(x)\,dx$$

следует ограниченность функции F в некоторой окрестности $U(+\infty) = \{x : x > b\}$ бесконечно удаленной точки $+\infty$ (см. свойство 1^0 в п. 5.10). На отрезке же $[a, b]$ функция F ограничена, ибо она непрерывна. В результате функция F ограничена на всей полупрямой $x \geqslant a$. Функция F является первообразной функции f; тем самым функция f имеет ограниченную первообразную при $x \geqslant a$.

Таким образом, для интеграла $\int_a^{+\infty} f(x)[g(x) - c]\,dx$ выполнены все условия признака Дирихле, поэтому этот интеграл сходится. В силу доказанного, из равенства (29.37) следует сходимость интеграла $\int_a^{+\infty} f(x)g(x)\,dx$.

Замечание 2. Усовершенствовав доказательства теорем 5 и 6, можно показать, что признаки Дирихле и Абеля сходимости интегралов остаются справедливыми, если у функции f условие ее непрерывности заменить условием ее интегрируемости на любом конечном отрезке $[a, b]$, а у функции g отбросить требование ее непрерывной дифференцируемости, оставив все остальные.

Примеры. 1. Применим признак Дирихле к исследованию сходимости интеграла

$$\int_a^{+\infty} \frac{\sin x}{x^\alpha}\, dx, \quad \alpha > 0. \tag{29.38}$$

Функция $f(x) = \sin x$ имеет ограниченную первообразную $F(x) = -\cos x$, а непрерывно дифференцируемая функция $g(x) = 1/x_\alpha$ при $\alpha > 0$ монотонно убывает и стремится к нулю при $x \to +\infty$. Все условия теоремы 5 выполнены, поэтому интеграл (29.38) сходится.

2. Следует, однако, иметь в виду, что признак Дирихле дает только достаточные, а не необходимые условия сходимости интеграла, поэтому не всегда с его помощью можно решить вопрос о сходимости интеграла. Например, исследуем сходимость интеграла

$$\int_1^{+\infty} \frac{\sin x\, dx}{x^\alpha - \sin x}, \quad \alpha > 0. \tag{29.39}$$

Попытаемся применить признак Дирихле, положив $f(x) = \sin x$ и $g(x) = \dfrac{1}{x^\alpha - \sin x}$. Очевидно, что $g(x) \to 0$ при $x \to \infty$.

Найдем производную:

$$g'(x) = \frac{\alpha x^{\alpha - 1} - \cos x}{(x^\alpha - \sin x)^2}.$$

Отсюда видно, что при $\alpha < 1$ эта производная при $x \to +\infty$ бесконечно много раз меняет знак и, следовательно, сама функция $g(x)$ не является монотонно убывающей.

Таким образом, при $\alpha < 1$ признак Дирихле неприменим указанным способом к выяснению вопроса о сходимости интеграла (29.39). В этом случае естественно попробовать прибегнуть снова к методу выделения главной части.

Применяя разложение функции $(1 - t)^{-1}$, $-1 < t < 1$, по формуле Тейлора (см. п. 13.3), при $x \to \infty$ получим

$$\frac{\sin x}{x^\alpha - \sin x} = \frac{\sin x}{x^\alpha} \frac{1}{1 - \dfrac{\sin x}{x^\alpha}} =$$

$$= \frac{\sin x}{x^\alpha}\left[1 + \frac{\sin x}{x^\alpha} + o\left(\frac{1}{x^\alpha}\right)\right] = \frac{\sin x}{x^\alpha} + \frac{\sin^2 x}{x^{2\alpha}} + o\left(\frac{1}{x^{2\alpha}}\right) =$$

$$= \frac{\sin x}{x^\alpha} + \frac{1}{2x^{2\alpha}} - \frac{\cos 2x}{2x^{2\alpha}} + o\left(\frac{1}{x^{2\alpha}}\right). \tag{29.40}$$

Интегралы

$$\int_1^{+\infty} \frac{\sin x}{x^\alpha}\,dx \quad \text{и} \quad \int_1^{\infty} \frac{\cos 2x}{x^{2\alpha}}\,dx \qquad (29.41)$$

сходятся по признаку Дирихле при всех $\alpha > 0$. Интеграл же

$$\int_1^{+\infty} \left[\frac{1}{2x^{2\alpha}} + o\left(\frac{1}{x^{2\alpha}}\right)\right] dx \qquad (29.42)$$

сходится при $2\alpha > 1$, т. е. при $\alpha > 1/2$, и расходится при $\alpha \leqslant 1/2$. Действительно, из формулы (29.40) вытекает, что функция $o\left(\frac{1}{x^{2\alpha}}\right)$ в указанной формуле непрерывна по x при $x \geqslant 1$, $\alpha > 0$, и, следовательно, имеет смысл говорить об интеграле (29.42). Функции $\frac{1}{x^{2\alpha}}$ и $\frac{1}{x^{2\alpha}} + o\left(\frac{1}{x^{2\alpha}}\right)$ неотрицательны в некоторой окрестности $+\infty$ и эквивалентны при $x \to +\infty$, поэтому интеграл (29.42) сходится и расходится при тех же значениях параметра α, что и интеграл $\int_1^{+\infty} \frac{dx}{x^{2\alpha}}$ (см. следствие из теоремы 1 в п. 29.3).

Таким образом, при $\alpha > 1/2$ все интегралы (29.41) и (29.42) сходятся, значит, в силу (29.40), сходится и интеграл (29.39). При $0 < \alpha \leqslant 1/2$ интегралы (29.41) сходятся, а интеграл (29.42) расходится, следовательно, расходится и интеграл (29.39).

Заметим, что при $\alpha \leqslant 0$ интеграл (29.39) расходится. Действительно, в этом случае знаменатель подынтегральной функции обращается в нуль бесконечно много раз; причем если $x_0^\alpha - \sin x_0 = 0$, то функция $x^\alpha - \sin x$ в окрестности точки x_0, согласно формуле Тейлора, имеет (почему?) вид $x^\alpha - \sin x = (x - x_0)^k \varphi(x)$, где k — некоторое натуральное число, а $\varphi(x_0) \neq 0$. По условию, $\sin x_0 = x_0^\alpha \neq 0$, поэтому в каждой такой точке x_0 функция $\frac{\sin x}{x^\alpha - \sin x}$ имеет неинтегрируемую особенность.

Следует обратить внимание на то, что для каждого фиксированного $\alpha > 0$ функции $\frac{\sin x}{x^\alpha - \sin x}$ и $\frac{\sin x}{x^\alpha}$ эквивалентны при $x \to +\infty$, т. е.

$$\frac{\sin x}{x^\alpha} = \varepsilon(x)\frac{\sin x}{x^\alpha - \sin x},$$

где $\varepsilon(x) = 1 - \frac{\sin x}{x^\alpha} \to 1$ при $x \to +\infty$, однако если $0 < \alpha \leqslant \frac{1}{2}$, то интеграл (29.39) от первой из них расходится, а интеграл (29.38) от второй сходится.

Таким образом, замена подынтегральной функции на эквивалентную может изменить сходимость интеграла (если, конечно, интеграл не сходится абсолютно).

3. Исследуем на сходимость и абсолютную сходимость интеграл

$$\int_1^{+\infty} \operatorname{tg}\left(\frac{\sin x}{x}\right) dx. \tag{29.43}$$

Так как $\left|\operatorname{tg}\left(\frac{\sin x}{x}\right)\right| \sim \frac{|\sin x|}{x}$ при $x \to +\infty$ и интеграл $\int_1^{+\infty} \frac{|\sin x|}{x}\, dx$ расходится (см. (29.32)), то расходится и интеграл $\int_1^{+\infty} \left|\operatorname{tg}\left(\frac{\sin x}{x}\right)\right| dx$, т. е. интеграл (29.43) не сходится абсолютно.

Легко проверить, что при $y \to 0$

$$\operatorname{tg} y = y + O(y^3), \tag{29.44}$$

причем в качестве окрестности в определении символа O (см. определение 1 в п. 8.2), здесь можно взять интервал $(-1, 1)$: существует такая постоянная $c > 0$, что $|O(y^3)| \leqslant c|y^3|$, $|y| < 1$.

Далее, в силу формулы (29.44) при $y = \frac{\sin x}{x}$, интеграл (29.43) можно представить в виде

$$\int_1^{+\infty} \operatorname{tg}\left(\frac{\sin x}{x}\right) dx = \int_1^{+\infty} \frac{\sin x}{x}\, dx + \int_1^{+\infty} O\left(\frac{1}{x^3}\right) dx. \tag{29.45}$$

Интеграл $\int_1^{+\infty} \frac{\sin x}{x}\, dx$ сходится (например, по признаку Дирихле), а интеграл $\int_1^{+\infty} O\left(\frac{1}{x^3}\right) dx$ абсолютно сходится, поэтому интеграл (29.43) — сходящийся.

4. Интеграл $\int_0^{+\infty} \frac{\sin x \operatorname{arctg} x}{x^\alpha}\, dx$, $\alpha > 0$, в силу признака Абеля, сходится.

В самом деле, как мы уже знаем (см. пример 1), интеграл $\int_0^{+\infty} \frac{\sin x}{x^\alpha}\, dx$ сходится, а функция $g(x) = \operatorname{arctg} x$ ограничена и монотонна.

УПРАЖНЕНИЕ 18. Исследовать на сходимость в абсолютную сходимость интеграл $\int\limits_0^{+\infty} \frac{\sin x dx}{(x + \cos x)^{\alpha}}$, $-\infty < \alpha < +\infty$.

29.7. Асимптотическое поведение интегралов с переменными пределами интегрирования

Часто при решении задач оказывается необходимым не только установить сходимость или расходимость рассматриваемого интеграла, но и уметь оценить в определенном смысле порядок «скорости» его сходимости или характер расходимости. Мы не будем здесь доказывать общих теорем, относящихся к этому вопросу (о некоторых общих методах изучения асимптотического поведения функций см. п. 33.10*), а лишь проиллюстрируем его на отдельных примерах отыскания порядка интегралов с переменным верхним пределом, когда они либо стремятся к нулю, либо к бесконечности. Именно, если, например, интеграл $\int\limits_a^{+\infty} f(x)\, dt$ сходится, то будет исследоваться порядок стремления к нулю при $x \to +\infty$ интеграла $\int\limits_x^{+\infty} f(x)\, dt$, т. е. отыскиваться явно заданная функция, эквивалентная при $x \to +\infty$ этому интегралу. Эта функция и называется *скоростью сходимости сходящегося несобственного интеграла* $\int\limits_a^{+\infty} f(x)\, dt$. Если же интеграл $\int\limits_a^{+\infty} f(t)\, dt$ расходится и равен $+\infty$ или $-\infty$, то будет изучаться порядок стремления к бесконечности при $x \to +\infty$ интеграла $\int\limits_a^{x} f(t)\, dt$, т. е. отыскиваться явно заданная функция, эквивалентная при $x \to +\infty$ теперь уже интегралу $\int\limits_a^{x} f(t)\, at$. Эта функция называется *скоростью расходимости расходящегося несобственного интеграла*.

Примеры. 1. Исследуем интеграл

$$\int\limits_2^{+\infty} \frac{dt}{l^{\alpha + 1} \ln^{\beta} t} \tag{29.46}$$

при различных действительных значениях параметров α и β. Рассмотрим сначала случай $\alpha > 0$ и любого $\beta \in \boldsymbol{R}$. При

таких значениях параметров интеграл (29.46) сходится, что легко устанавливается по признаку сравнения, если в качестве функции сравнения взять, например, функцию $g(t) = t^{-\alpha/2-1}$, интеграл от которой $\int\limits_2^{+\infty}\frac{dt}{t^{\alpha/2+1}}$ сходится.

В силу сходимости интеграла (29.46), при указанных значениях параметров α и β второе слагаемое правой части равенства $\int\limits_2^{+\infty}\frac{dt}{t^{\alpha+1}\ln^\beta t} = \int\limits_2^{x}\frac{dt}{t^{\alpha+1}\ln^\beta t} + \int\limits_x^{+\infty}\frac{dt}{t^{\alpha+1}\ln^\beta t}$ стремится к 0 при $x \to \infty$. Изучим порядок его убывания, а именно покажем справедливость асимптотического равенства

$$\int\limits_x^{+\infty}\frac{dt}{t^{\alpha+1}\ln^\beta t} \sim \frac{1}{\alpha x^\alpha \ln^\beta x},\ x \to +\infty. \qquad (29.47)$$

Для доказательства положим

$$F(x) \overset{\text{def}}{=} \int\limits_x^{+\infty}\frac{dt}{l^{\alpha+1}\ln^\beta t}, \quad \Phi(x) \overset{\text{def}}{=} \frac{1}{\alpha x^\alpha \ln^\beta x}.$$

В силу сходимости интеграла (29.46), при $\alpha > 0$, $\beta \in \boldsymbol{R}$ имеем $\lim\limits_{x\to+\infty} F(x) = 0$. Очевидно, и $\lim\limits_{x\to+\infty}\Phi(x) = 0$. Так как

$$F'(x) = -\frac{1}{x^{\alpha+1}\ln^\beta x}, \quad \Phi'(x) = -\frac{1}{x^{\alpha+1}\ln^\beta x} - \frac{\beta}{\alpha x^{\alpha+1}\ln^{\beta+1} x},$$

то, применив правило Лопиталя, получим

$$\lim_{x\to+\infty}\frac{\Phi(x)}{F(x)} = \lim_{x\to+\infty}\frac{\Phi'(x)}{F'(x)} = \lim_{x\to+\infty}\left(1 + \frac{\beta}{\alpha \ln x}\right) = 1,$$

т. е. соотношение (29.47) доказано.

В случае $\alpha = 0$, $\beta > 1$ непосредственным интегрированием получим даже явное выражение интересующего нас интеграла

$$\int\limits_x^{+\infty}\frac{dt}{t\ln^\beta t} = \int\limits_x^{+\infty}\frac{d\ln t}{\ln^\beta t} = \frac{1}{(1-\beta)\ln^{\beta-1} t}\Bigg|_x^{+\infty} = \frac{1}{(\beta-1)\ln^{\beta-1} t}. \qquad (29.48)$$

Покажем теперь, что для $\alpha < 0$ и любого $\beta \in \boldsymbol{R}$ интеграл (29.46) расходится. Чтобы убедиться в этом, достаточно снова в качестве функции сравнения взять функцию $g(t) = t^{-\frac{\alpha}{2}-1}$, где в данном случае, т. е. при $\alpha < 0$, интеграл $\int\limits_2^{+\infty} g(x)\,dt$ расходится.

Докажем, что при $\alpha < 0$ и любом $\beta \in \boldsymbol{R}$ имеет место асимптотическое равенство

$$\int_2^x \frac{dt}{t^{\alpha+1}\ln^\beta t} \sim -\frac{1}{\alpha x^\alpha \ln^\beta x}, \quad x \to +\infty.$$

Положим

$$F(x) \stackrel{\text{def}}{=} \int_2^x \frac{dt}{t^{\alpha+1}\ln^\beta t}, \quad \Phi(x) = -\frac{1}{\alpha x^\alpha \ln^\beta x},$$

тогда $\lim\limits_{x \to +\infty} \mathrm{F}(x) = \lim\limits_{x \to +\infty} \Phi(x) = +\infty$. Найдем производные функции $F(x)$ и $\Phi(x)$:

$$F'(x) = \frac{1}{x^{\alpha+1}\ln^\beta x}, \quad \Phi'(x) = \frac{1}{x^{\alpha+1}\ln^\beta x} + \frac{\beta}{\alpha x^{\alpha+1}\ln^{\beta+1} x}.$$

Применив правило Лопиталя, получим

$$\lim_{x \to +\infty} \frac{\Phi(x)}{F(x)} = \lim_{x \to +\infty} \frac{\Phi'(x)}{F'(x)} = \lim_{x \to +\infty} \left(1 + \frac{\beta}{\alpha \ln x}\right) = 1,$$

т. е. равенство (29.49) доказано.

Для оставшихся значений параметров α и β интеграл

$$\int_2^x \frac{dt}{t^\alpha \ln^\beta x} \tag{29.50}$$

вычисляется в элементарных функциях. Если $\alpha = 0$ и $\beta < 1$, то

$$\int_2^x \frac{dt}{t\ln^\beta t} = \int_2^x \frac{d\ln t}{\ln^\beta t} = \left.\frac{1}{(1-\beta)\ln^{\beta-1} t}\right|_2^x = \frac{\ln^{1-\beta} x - \ln^{1-\beta} 2}{1-\beta},$$

а если $\alpha = 0$, $\beta = 1$, то

$$\int_2^x \frac{dt}{t\ln t} = \int_2^x \frac{d\ln t}{\ln t} = \ln \ln t\Big|_2^x = \ln \frac{\ln x}{\ln 2}.$$

В обоих случаях

$$\lim_{x \to +\infty} \int_2^x \frac{dt}{t\ln^\beta x} = +\infty,$$

поэтому интеграл (29.46) при $\alpha = 0$ и $\beta \leqslant 1$ расходится.

Итак, интеграл (29.46) сходится при $\alpha > 0$ и любом $\beta \in \boldsymbol{R}$, а также при $\alpha = 0$ и $\beta > 1$; при этом установлены асимптотическое, и соответственно точное равенства (29.47) и (29.48) для интеграла $\int_x^{+\infty} \frac{dt}{l^{\alpha+1}\ln^\beta t}$. При остальных значениях параметров α и β интеграл (29.46) расходится и получена асимптотическая или точная характеристика интеграла (29.50).

2. Рассмотрим интеграл

$$\int_T^{+\infty} \frac{f(t)}{t}\,dt, \tag{29.51}$$

где при $t \geqslant T$ функция f непрерывна, неотрицательна

$$f(t) \geqslant 0, \tag{29.52}$$

и для любого $t \geqslant T$ выполняется равенство

$$f(t + T) = f(t). \tag{29.53}$$

Будем функцию f в этом случае для краткости называть также периодической с перидом T (ср. с определением периодической функции в п. 24.1, пример 2). Пусть еще интеграл от функции f по периоду положителен

$$\int_T^{2T} f(t)\,dt > 0. \tag{29.54}$$

Покажем, что интеграл (29.51) расходится и что

$$\int_T^x \frac{f(t)}{t}\,dt \asymp \ln x,\ \ x \to +\infty, \tag{29.55}$$

т. е. что функция в левой части этой формулы при $x \to +\infty$ имеет порядок $\ln x$ (см. п. 8.2).

Функция f непрерывна, поэтому она ограничена на отрезке $[t,\ 2T]$ и, следовательно, в силу периодичности, ограничена при $t \geqslant T$, т. е. существует такое число $M > 0$, что и для всех $t \geqslant T$ выполняется неравенство

$$f(t) \leqslant M. \tag{29.56}$$

В силу этого, имеем

$$\int_T^x \frac{f(t)}{t}\,dt \underset{(29.56)}{\leqslant} M\int_T^x \frac{dt}{t} = \ln x - \ln T = O(\ln x),\ \ x \geqslant T. \tag{29.57}$$

Обозначим теперь через J интеграл от функции f по периоду, т. е.

$$J = \int_T^{2T} f(t)\,dt. \tag{29.58}$$

Произведя в записанных ниже интегралах замену переменного $t = u + (k - 1)T$ на отрезках $[kT,\ (k + 1)T]$, $k = 2, 3, \ldots$, получим

$$\int_T^{nT} \frac{f(t)}{t}\,dt = \sum_{k=1}^{n-1} \int_{kT}^{(k+1)T} \frac{f(t)}{t}\,dt =$$

$$= \sum_{k=1}^{n-1} \int_T^{2T} \frac{f(u + (k-1)T)}{u + (k-1)T}\,du \underset{(29.53)}{=} \frac{1}{T}\sum_{k=1}^{n-1} \int_T^{2T} \frac{f(u)}{\frac{u}{T} + k - 1}\,du \geqslant$$

$$\geqslant \frac{1}{T}\sum_{k=1}^{n-1} \frac{1}{2 + k - 1} \int_T^{2T} f(u)\,du \underset{(29.58)}{\geqslant} \frac{J}{T}\sum_{k=1}^{n-1} \frac{1}{k+1}. \tag{29.59}$$

Заметим, что для чисел x, лежащих на отрезке $[k+1, k+2]$, выполняется неравенство $\frac{1}{x} \leqslant \frac{1}{k+1}$, интегрированием которого получается неравенство

$$\int_{k+1}^{k+2} \frac{dx}{x} \leqslant \frac{1}{k+1}. \qquad (29.60)$$

Поэтому

$$\sum_{k=1}^{n-1} \frac{1}{k+1} \underset{(29.60)}{\geqslant} \sum_{k=1}^{n-1} \int_{k+1}^{k+2} \frac{dx}{x} = \int_{2}^{n+1} \frac{dx}{x} =$$
$$= \ln(\mathrm{n}+1) - \ln 2 \geqslant \ln(n+1).$$

Подставив эту оценку в неравенство (29.59), получим

$$\int_{T}^{nT} \frac{f(t)}{t}\, dt \geqslant \frac{J}{T} \ln(n+1). \qquad (29.61)$$

Заметим, что

$$\lim_{n \to \infty} \frac{\ln(n+1)}{\ln(n+1)T} = \lim_{n \to \infty} \frac{1}{1 + \frac{\ln T}{\ln(n+1)}} = 1,$$

поэтому существует такое натуральное n_0, что при $n > n_0$ выполняется неравенство

$$\frac{\ln(n+1)}{\ln(n+1)T} \geqslant \frac{1}{2}. \qquad (29.62)$$

Далее, для каждого числа x существует такое целое n, что

$$nT \leqslant x < (n+1)T. \qquad (29.63)$$

Теперь для любого x, для которого в неравенстве (29.63) имеет место неравенство $n \geqslant n_0$, получаем

$$\int_{0}^{x} \frac{f(t)}{t}\, dt \underset{(29.63)}{\geqslant} \int_{0}^{nT} \frac{\sin^2 t}{t}\, dt \underset{(29.61)}{\geqslant} \frac{J}{T} \ln(n+1) \underset{(29.62)}{\geqslant}$$
$$\underset{(29.62)}{\geqslant} \frac{J}{2T} \ln(n+1)T \underset{(29.63)}{\geqslant} \frac{J}{2T} \ln x. \qquad (29.64)$$

Неравенства (29.57) и (29.64) и доказывают справедливость формулы (29.55).

В формуле (29.55) в качестве функции f можно взять любую конкретную функцию, удовлетворяющую перечисленным выше условиям, при этом, в силу доказанного, соответ-

ствующие интегралы всегда будут иметь порядок $\ln x$. Например,

$$\int_1^x \frac{\ln(\cos t + 2)}{t}\, dt \asymp \ln x,\ x \to +\infty,$$

$$\int_0^x \frac{\sin^2 t}{t}\, dt \asymp \ln x,\ x \to +\infty.$$

Однако иногда удается получить более точную оценку. Так, для второго из интегралов имеем

$$\int_0^x \frac{\sin^2 t}{t}\, dt = \int_0^1 \frac{\sin^2 t}{t}\, dt + \int_1^x \frac{1-\cos 2t}{2t}\, dt =$$

$$= \int_0^1 \frac{\sin^2 t}{t}\, dt + \frac{1}{2}\int_1^x \frac{dt}{t} - \frac{1}{2}\int_1^x \frac{\cos 2t}{t}\, dt =$$

$$= \frac{1}{2}\ln x + \int_0^1 \frac{\sin^2 t}{t}\, dt - \frac{1}{2}\int_1^x \frac{\cos 2t}{t}\, dt. \qquad (29.65)$$

Так как $\lim\limits_{t\to 0} \frac{\sin^2 t}{t} = 0$, то функция $\frac{\sin^2 t}{t}$, доопределенная нулем при $t = 0$, будет непрерывна и, следовательно, интегрируема на отрезке $[0, 1]$, т. е. $\int_0^1 \frac{\sin^2 t}{t}\, dt$ конечен. Интеграл $\int_1^{+\infty} \frac{\cos 2t}{t}\, dt$ сходится (это, например, сразу следует из признака Дирихле, см. п. 29.6). Из сказанного вытекает, что функция

$$F(x) \stackrel{\text{def}}{=} \int_0^1 \frac{\sin^2 t}{t}\, dt - \frac{1}{2}\int_1^x \frac{\cos 2t}{t}\, dt, \qquad (29.66)$$

будучи непрерывной для всех $x \geqslant 0$ и имея конечный предел

$$\lim_{x\to+\infty} F(x) = \int_0^1 \frac{\sin^2 t}{t}\, dt - \frac{1}{2}\int_1^{+\infty} \frac{\cos 2t}{t}\, dt,$$

ограничена на неотрицательной полуоси. Поэтому из равенства

$$\int_0^x \frac{\sin^2 t}{t}\, dt \underset{\substack{(29.65)\\(29.66)}}{=} \frac{1}{2}\ln x + F(x)$$

явствует, что

$$\int_0^x \frac{\sin^2 t}{t}\, dt \sim \frac{1}{2}\ln x, \quad x \to +\infty,$$

т. е. в этом случае удается определить не только порядок интеграла с переменным пределом интегрирования x, но и его

асимптотическое поведение при $x \to +\infty$; интеграл эквивалентен $\frac{1}{2}\ln x$.

В рассмотренных примерах асимптотическое поведение интегралов установлено с помощью более или менее специальных методов, оказавшихся удобными в рассмотренных конкретных случаях. Более общим методом, часто дающим возможность находить асимптотическое поведение интегралов, является обычное интегрирование по частям.

3. Рассмотрим в качестве примера *интегралы Френеля*[1]

$$\int_0^{\infty} \cos\theta^2\, d\theta,\ \int_0^{\infty} \sin\theta^2\, d\theta,$$

скорость сходимости которых определяется порядком убывания интегралов

$$\int_x^{+\infty} \cos\theta^2\, d\theta,\ \int_x^{+\infty} \sin^2 d\theta,\ x > 0. \qquad (29.67)$$

Изучение асимптотического поведения интегралов (29.67) при $x \to +\infty$ проводится одинаковым методом. Поэтому рассмотрим только один из них, например первый. Сделав в нем замену переменной $\theta^2 = t$, сразу убеждаемся по признаку Дирихле, что он сходится. Затем, дважды проинтегрировав по частям получившийся интеграл, будем иметь

$$\int_x^{+\infty} \cos\theta^2\, d\theta = \frac{1}{2}\int_{x^2}^{+\infty} \frac{\cos t}{\sqrt{t}}\, dt =$$

$$= \frac{1}{2}\frac{\sin t}{\sqrt{t}}\Bigg|_{x^2}^{+\infty} + \frac{1}{4}\int_{x^2}^{+\infty} \frac{\sin t}{t\sqrt{t}}\, dt = -\frac{\sin^2 x}{2x} + \frac{1}{4}\int_{x^2}^{+\infty} \frac{\sin t}{t\sqrt{t}}\, dt =$$

$$= -\frac{\sin x^2}{2x} + \frac{\cos x^2}{4x^3} - \frac{3}{8}\int_{x^2}^{+\infty} \frac{\cos t}{t^2\sqrt{t}}\, dt \qquad (29.68)$$

(согласно прежней терминологии (см. п. 29.5), мы с помощью интегрирования по частям улучшили сходимость интеграла).

Поскольку $\frac{\cos x^2}{4x^3} = O\left(\frac{1}{x^3}\right)$, $x \to \infty$, и

$$\left|\int_{x^2}^{+\infty} \frac{\cos t}{t^2\sqrt{t}}\, dt\right| \leqslant \int_x^{+\infty} \frac{dt}{t^2\sqrt{t}} = -\frac{2}{3t^{3/2}}\Bigg|_{x^2}^{+\infty} = \frac{2}{3x^3},$$

[1] А. Френель (1788—1827) — французский физик.

будем иметь

$$\int_{x^2}^{+\infty} \frac{\cos t}{t^2\sqrt{t}}\,dt = O\left(\frac{1}{x^3}\right),\quad x \to +\infty.$$

Следовательно,

$$\int_{x}^{+\infty} \cos\theta^2\,d\theta = -\frac{\sin x^2}{2x} + O\left(\frac{1}{x^3}\right),\quad x \to +\infty.$$

Таким образом, нам удалось с точностью до $O\left(\frac{1}{x^3}\right)$, $x \to +\infty$, найти простое выражение для интеграла $\int_{x}^{+\infty} \cos\theta^2\,d\theta$, дающее, в частности, представление о характере его убывания при $x \to +\infty$. Если произвести дальнейшее интегрирование по частям интеграла в правой части формулы (29.68), то можно получить асимптотические формулы для интеграла $\int_{x}^{+\infty} \cos\theta^2\,d\theta$ с точностью до $O\left(\frac{1}{x^{2n+1}}\right)$, $x \to +\infty$, при любом натуральном n.

Аналогично рассматривается случай интеграла $\int_{x}^{+\infty} \sin\theta^2\,d\theta$.

Предметно-именной указатель

Указатель основных обозначений

$\exists x$ — существует такое x
$\forall x$ — для любого x
$A \Rightarrow B$ — из высказывания A следует высказывание B
$A \Leftrightarrow B$ — высказывание A равносильно высказыванию B
$\overset{\text{def}}{=}$ — утверждение справедливо по определению
$\sum_{k=1}^{n} a_k$ — сумма n слагаемых $a_1, \ldots, a_n$
$\square$ — знак окончания доказательства
$a \in A$, $A \ni a$ — элемент a принадлежит множеству A
$a \notin A$, $A \not\ni a$ — элемент a не принадлежит множеству A
$A = \{a, b, c, \ldots\}$ — множество A состоит из элементов $a, b, c\ldots$
$A = \{a\colon \ldots\}$ — множество A состоит из элементов, обладающих определенными свойствами (указанными после двоеточия)
$\varnothing$ — пустое множество
$A \subset B$, $B \supset A$ — множество A содержится в множестве B
$A \cup B$ — объединение множеств A и B
$A \cap B$ — пересечение множеств A и B
$A \backslash B$ — разность множеств A и B
$\boldsymbol{N}$ — множество натуральных чисел
$\boldsymbol{Z}$ — множество целых чисел
$\boldsymbol{Q}$ — множество рациональных чисел
$\boldsymbol{I}$ — множество иррациональных чисел
$\boldsymbol{R}$ — множество действительных чисел, числовая прямая
$\overline{\boldsymbol{R}}$ — расширенное множество действительных чисел, расширенная числовая прямая
$\boldsymbol{C}$ — множество комплексных чисел
i — мнимая единица
$\operatorname{Re} z$ — действительная часть комплексного числа z
$\operatorname{Im} z$ — мнимая часть комплексного числа z
$|z|$ — модуль комплексного числа z
$\operatorname{Arg} z$ — аргумент комплексного числа z
$\arg z$ — значение аргумента φ комплексного числа z, удовлетворяющее неравенству $-\pi < \varphi \leqslant \pi$
$\bar{z}$ — число, сопряженное комплексному числу z
$[a, b]$ — отрезок
(a, b) — интервал
$[a, b)$, $(a, b]$ — полуинтервалы
A_n^k — размещения из n элементов по k элементов
P_n — перестановки из n элементов
C_n^k — сочетания из n элементов по k элементов

A/B — сечение множества действительных или рациональных чисел, образованное множествами A и B

$\sup X$, $\sup\limits_{x \in X} \{x\}$ — верхняя грань множества X

$\inf X$, $\inf\limits_{x \in X} \{x\}$ — нижняя грань множества X

$\overline{X}$ — замыкание множества X

$\{x_n\}$; x_n, $n = 1, 2, \ldots$, — последовательность элементов x_n

$\lim\limits_{n \to \infty} x_n = a$ — число a есть предел последовательности $\{x_n\}$

$\overline{\lim\limits_{n \to \infty}}\, x_n$ — верхний предел последовательности $\{x_n\}$

$\underline{\lim\limits_{n \to \infty}}\, x_n$ — нижний предел последовательности $\{x_n\}$

$f: X \to Y$ — отображение множества X в множество Y

$y = f(x)$ — переменная y является функцией переменной x

$f(x_0)$, $f(x)|_{x = x_0}$ — значение функции $f(x)$ в точке x_0

$f|_A$, f_A — сужение функции f на множество A

$g \circ f$ — композиция (суперпозиция) функций f и g

f^{-1} — обратная функция для функции f

e — основание натуральных логарифмов

$\ln x$ — натуральный логарифм числа x

$\operatorname{sh} x$ — гиперболический синус

$\operatorname{ch} x$ — гиперболический косинус

$\operatorname{th} x$ — гиперболический тангенс

$\operatorname{cth} x$ — гиперболический котангенс

$\operatorname{Arsh} x$ — ареасинус

$\operatorname{Arch} x$ — ареакосинус

$\operatorname{sign} x$ — функция «знак x»

$[x]$ — функция «целая часть числа x»

$\sup f$, $\sup\limits_X f$, $\sup\limits_{x \in X} f(x)$ — верхняя грань функции f на множестве X

$\inf f$, $\inf\limits_X f$, $\inf\limits_{x \in X} f(x)$ — нижняя грань функции f на множестве X

$\max f$, $\max\limits_X f$ — наибольшее значение функции f на множестве X

$\min f$, $\min\limits_X f$ — наименьшее значение функции f на множестве X

$\lim\limits_{x \to x_0} f(x) = a$, $f(x) \to a$ при $x \to x_0$ — число a есть предел функции $f(x)$ в точке x_0

$\lim\limits_{\substack{x \to x_0 \\ x \in E}} f(x) = a$ — число a есть предел функции $f(x)$ по множеству E в точке x_0

$\lim f(x_0 - 0)$, $\lim\limits_{x \to x_0 - 0} f(x)$ — предел слева функции $f(x)$ в точке x_0, $x \neq x_0$

$\lim f(x_0 + 0)$, $\lim\limits_{x \to x_0 + 0} f(x)$ — предел справа функции $f(x)$ в точке x_0, $x \neq x_0$

$U(x_0, \varepsilon)$ —ε-окрестность точки x_0

$\mathring{U}(x_0, \varepsilon)$ — проколотая ε-окрестность точки x_0

$X_{\leqslant}(x_0)$ — пересечение множества X с лучом $x \leqslant x_0$

$X_{\geqslant}(x_0)$ — пересечение множества X с лучом $x \geqslant x_0$

$f(x) = O(g(x))$ при $x \to x_0$ — функция f ограничена по сравнению с функцией g при $x \to x_0$

$f(x) \asymp g(x)$ при $x \to x_0$ — функция f имеет тот же порядок, что и функция g, при $x \to x_0$

$f(x) \sim g(x)$ при $x \to x_0$ — функции f и g эквивалентны (асимптотически равны) при $x \to x_0$

$\alpha(x) = o(f(x))$ при $x \to x_0$ — функция α есть бесконечно малая по сравнению с функцией f при $x \to x_0$

Δx — приращение аргумента функции $y = f(x)$

Δy — приращение функции функции $y = f(x)$

$f(x)$, y', y'_x, $\frac{dy}{dx}$ — производная функции $y = f(x)$

$f'_+(x)$ — правая производная функции $f(x)$ в точке x

$f'_-(x)$ — левая производная функции $f(x)$ в точке x

dy, $df(x)$ — дифференциал функции $y = f(x)$

$f''(x)$, $f^{(2)}(x)$, y'' — вторая производная функции $y = f(x)$

$f^{(n)}(x)$, $y^{(n)}$ — производная n-го порядка функции $y = f(x)$

$d^n y$ — дифференциал n-го порядка функции $y = f(x)$

$\boldsymbol{r}(t)$ — векторная функция

$\boldsymbol{ab}$, $(\boldsymbol{a}, \boldsymbol{b})$ — скалярное произведение векторов $\boldsymbol{a}$ и $\boldsymbol{b}$

$\boldsymbol{a} \times \boldsymbol{b}$, $[\boldsymbol{a}, \boldsymbol{b}]$ — векторное произведение векторов $\boldsymbol{a}$ и $\boldsymbol{b}$

$\widehat{\boldsymbol{ab}}$ — угол между векторами $\boldsymbol{a}$ и $\boldsymbol{b}$

$\int f(x)\,dx$ — неопределенный интеграл от функции $f(x)$

$\int_b^a f(x)\,dx$ — определенный интеграл от функции $f(x)$ на отрезке $[a, b]$; несобственный интеграл по промежутку с концами a и b, $-\infty \leqslant a < b \leqslant +\infty$

$\Phi(x)\Big|_a^b$, $[\Phi(x)]_a^b$ — разность $\Phi(b) - \Phi(a)$

mes G — мера (площадь, объем) открытого множества

Оглавление

Глава 1

Дифференциальное исчисление функций одной переменной

Глава 2
Интегральное исчисление функций одной переменной